The Binary Black Holes of LIGO and Virgo

Javier Roulet

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Physics
Adviser: Matias Zaldarriaga

September 2021

Abstract

Over the last six years, the LIGO and Virgo gravitational wave detectors have revolutionized gravitational wave astronomy by discovering the first compact binary mergers. There is much to learn about how these systems form in nature, and these discoveries have allowed to start characterizing the astrophysical population of binary black holes. Many layers of data processing are needed in the path from raw gravitational wave data to inference of astrophysical implications. In this thesis, I worked on algorithms to search for signals from compact binary mergers, estimating their parameters and analyzing them collectively to infer properties of the astrophysical population, ultimately aimed at unveiling key questions in gravitational wave astronomy: What is the rate of binary mergers in the Universe? What is the distribution of masses and spins? What is the formation mechanism of merging binary black holes?

I construct a bank of waveform templates suitable for searching compact binary mergers in gravitational wave data through matched-filtering. The resulting bank is defined on a geometric space, whose notion of distance between waveforms naturally corresponds to their response mismatch. Beyond aiding intuition, this feature enables optimal placement of templates, dynamical refinement of the search, and powerful and robust signal quality tests. Using this template bank, my collaborators and I carried a search for binary black holes in public LIGO–Virgo data, confirming previous detections and identifying nine new events.

I compute the likelihood function for the parameters of the individual sources, such as black hole masses and spins. I derive a framework to combine these pieces of information into a likelihood for the collective distribution of these parameters, that accounts for measurement uncertainties, selection effects and statistical significance of the events. With this, I test and constrain phenomenological models for the distribution of binary black hole masses, spins, merger rates and cosmological evolution. I find that the mass distribution features a steep drop around $40\,\mathrm{M}_\odot$, as predicted by the pair-instability supernova mechanism; but also features an extended tail to higher masses. The distribution of spin orientations is anisotropic, disfavoring dynamical formation channels as the only pathway for merging binary black holes.

Acknowledgements

Over the years of my PhD, I have grown a lot as a person and as a scientist, and this has been largely due to the wonderful group of mentors, colleagues and friends around me.

First, thanks to Matias Zaldarriaga for being a most inspiring adviser. From you, I learned to tell not only which answers are right, but which questions are relevant, and which can hope to be answered. You have also given me lots of support and good advice. I would like to thank Barak Zackay, guru of statistics, algorithms, honesty and mind-boggling optimism in the face of almost-impossible projects; Teja Venumadhav, master of elegance and thoroughness; and Liang Dai, insightful and enthusiastic astrophysicist, for being mentors before collaborators. You have all made me a much better scientist. I am also grateful to David Radice for a year of attentive mentoring.

I would further like to thank Gabriel Mindlin for his trust and support early in my career; and Bill Bialek for the same and for kindly offering to take part in my thesis committee "to learn something new"—I hope this dissertation can live up to that expectation. I extend my thanks to Jo Dunkley and Lyman Page for agreeing to review this thesis.

Much of this work relies on access to LIGO–Virgo data. I thank the LIGO–Virgo Collaboration in general, and their Gravitational Wave Open Science Center in particular, for their open data policy.

My time in graduate school was quite enjoyable thanks to the good friends I made in Princeton and New York City. I am happy to thank Brandon, Justin, Chris, Seth, Diana, Fedor, Ho Tat, Xiaowen, Jorge, Mayu, Amel and Mandeep for that.

Finally, I would like to thank my parents and family for their love and support with all my projects, with a special mention to Paula for making 2020 the year we spent 5 months at Paula's. Above all, thanks to my wife Rocio for sharing this journey with me.

To my family.

Contents

Chapter 1

Introduction

1.1 Gravitational waves

Gravitational waves are disturbances of space-time that propagate at the speed of light, predicted over 100 years ago by General Relativity. Their physical effect is to stretch or compress distances between free-falling objects, along directions transverse to the propagation of the wave. Gravitational wave detectors operate on this principle to measure them. Currently there are three gravitational wave detectors in operation: two Advanced LIGO located in Hanford, WA, and Livingston, LA, in the United States, and Advanced Virgo in Cascina, Italy. Advanced LIGO and Advanced Virgo became operational in 2015 and 2017, respectively, and use laser interferometry to precisely monitor variations in the distance between test masses caused by gravitational waves.

Gravitational waves are produced in nature by rapidly accelerating, compact massive objects. A gravitational wave is characterized by the dimensionless strain tensor h_{ij}, which can be interpreted as the fractional change in distances it produces along the various directions. Emission of gravitational waves is described by the quadrupole formula (to leading order in the strength of the gravitational field and velocity of the source, and in the wave zone far from the source):

$$h_{ij} = \frac{2G}{c^4 D} \ddot{Q}_{ij}, \tag{1.1}$$

where G is the gravitational constant, c is the speed of light, D is the (luminosity) distance to the source and

$$Q_{ij} = \int \mathrm{d}^3x \, \rho(x) \left(x_i x_j - \frac{1}{3} \delta_{ij} r^2 \right) \tag{1.2}$$

is the traceless mass quadrupole moment, with ρ the mass density. The strain is highly suppressed by the factor $G/c^4 \sim 10^{-44}\,\mathrm{s^2 kg^{-1} m^{-1}}$ in Eq. (1.1), which makes detection of gravitational waves extremely challenging from a technological point of view. Detection is impossible for all but the loudest sources, which involve motion of objects multiple times heavier than the solar system at velocities comparable to the speed of light. The classic example of a source of gravitational waves is a binary system made of compact objects, like black holes or neutron stars, in tight orbit (and indeed these have been the only so far observed). Gravitational wave emission causes the binary to lose energy, tightening the orbit and further increasing gravitational wave emission. If the orbit is initially tight enough, this is a runaway process which ends in a violent merger of the two objects in less than the age of the universe. We can crudely estimate the order of magnitude of gravitational waves emitted from a binary as follows. For a binary with mass M and orbital separation R, the quadrupole moment is

$$Q \sim MR^2. \tag{1.3}$$

The timescale for its variation is given by the orbital frequency, $\ddot{Q} \sim \omega^2 Q$, which in turn follows from Newton's laws: $GM^2/R^2 \sim MR\omega^2$ or

$$\omega^2 \sim \frac{GM}{R^3}. \tag{1.4}$$

Putting everything together,

$$h \sim \frac{G}{c^4 D} MR^2 \frac{GM}{R^3} = \left(\frac{GM}{c^2}\right)^2 \frac{1}{RD} \sim \frac{R_S^2}{RD}, \tag{1.5}$$

where $R_S = 2GM/c^2$ is the Schwarzschild radius of a black hole of mass M. Emission is largest for small orbital separation R, which can only be as small as the component objects' size. This is why merging black holes or neutron stars, whose size is comparable to their Schwarzschild radius, are prime sources of gravitational waves. For these, $R \sim R_S$ and Eq. (1.5) becomes

$$h \sim \frac{R_S}{D} = 10^{-21} \frac{M}{\mathrm{M}_\odot} \frac{100\,\mathrm{Mpc}}{D}, \tag{1.6}$$

where $\mathrm{M}_\odot$ is the mass of the Sun and a megaparsec Mpc is comparable to intergalactic separations (as it turns out, it is necessary to probe out to hundreds of Mpc in order to observe events at a rate of several per year). This means that gravitational wave detectors need to be sensitive to strain on the order of 10^{-21}. For a kilometer-scale detector such as LIGO or Virgo, this requires measuring

a change in separation between the test masses of $\Delta L = hL \sim 10^{-18}\,\mathrm{m}$—a thousandth part of the diameter of a proton. Considering that test masses are made of atoms, and laser interferometry uses micron-wavelength light, this is an astonishing achievement. A historical review of how this formidable challenge was overcome is given in Saulson (2019).

Advanced LIGO achieved the first direct detection of gravitational waves from a binary black hole merger on 2015, September 14, which opened a new observational window to the universe and revolutionized astrophysics (Abbott et al., 2016a). Since then, three observing runs (O1, O2, O3) took place in 2015, 2017 and 2019, interspersed with periods of detector upgrades that progressively increased their sensitivity. Altogether, about fifty binary black holes and two binary neutron stars have been identified.

A compact binary merger signal exhibits a waveform of characteristic shape ("chirp"), with increasing frequency and amplitude during the inspiral and terminating at merger (Abbott et al., 2016d). By analyzing the detailed shape of the waveform and comparing it to physical models of binary mergers, we can estimate the parameters of the source, like masses, spins, location and orientation. This is very valuable to shed light on the astrophysics of these systems and their progenitors, particularly since binary black holes can only be observed through gravitational waves.

A pressing question that these observations can inform is what is the astrophysical origin of merging binary black holes. Several ideas have been put forward, broadly encompassing two classes of mechanisms: dynamical formation, where binaries are assembled from independently-formed black holes; or isolated formation, where a binary star evolves into a merging binary black hole. I briefly describe some of the main ideas below, see Mandel and Farmer (2018) for a review.

The leading mechanism to dynamically form a binary from unbound objects is through a three-body interaction that binds the binary and deposits the excess energy in the third object. Further few-body interactions are required to tighten the orbit of the binary, until gravitational wave emission can take over and lead to a merger within the age of the Universe. Dynamical formation can thus only happen in dense stellar clusters; where the rate of three-body encounters is appreciable. Such clusters include globular clusters (old clusters of $\sim 10^5$ stars, of which each galaxy typically has hundreds), nuclear clusters (dense clusters of $\sim 10^6$ stars at the center of most galaxies) or open clusters (small, young clusters of $\sim 10^3$ stars, where stars are born). A robust, testable prediction of dynamical formation channels is that the spin orientations of the independently-formed black holes in a merger are distributed isotropically, uncorrelated with each other and with the direction of the orbit. Dynamical formation channels also tend to produce binaries of black holes with comparable masses. This is because gravitational interactions induce mass-segregation: heavier objects move

quickly to the core of the cluster where they get paired with each other. Furthermore, three-body scatterings tend to eject the lighter object with high probability, which enhances exchange of companions for unequal-mass binaries. An interesting exception are "second-generation" mergers, that are possible within dynamic formation channels: if the remnant black hole is retained in the cluster after merger, it can become paired and merge again. Second-generation mergers would be characterized by a high mass, a $\sim 2:1$ mass ratio, and a large spin of the heavier black hole of $\approx 0.7\,GM^2/c$ due to the orbital angular momentum from the first merger.

In isolated-formation channels, instead, a binary star in a generic location of its galaxy (the "galactic field") evolves into a binary black hole that then merges. The main challenge in this pathway is that mechanisms other than gravitational waves are required to bring the objects together: a binary black hole with an initial separation larger than the maximum (giant phase) radius of its progenitor stars would take longer than the age of the Universe to merge by gravitational wave emission. One natural mechanism is a common-envelope phase, which is a short unstable mass transfer episode triggered by the expansion of one of the two stars. When a star exhausts hydrogen at its core, fusion temporarily stops and the core contracts and heats until helium fusion is ignited. The heating of the core causes the hydrogen envelope to expand; a star in a binary can overflow its Roche lobe (the region in the co-rotating frame in which trajectories are bound to that star), and transfer mass to its companion. Mass transfer has two effects: conservation of angular momentum requires the orbital separation to change (heavy-to-light transfer decreases the separation, and vice versa), it also alters the hydrodynamic equilibrium of the donor. Depending on the interplay between these effects, mass transfer can enhance or inhibit itself, leading to unstable or stable transfer, respectively. Unstable mass transfer triggers a common envelope episode that lasts only a few orbital periods; a large amount of orbital energy is dumped into the envelope of the star, which gets ejected. If this process stops before the two stars merge, it results in a tight binary that can ultimately merge by gravitational wave emission. After common envelope, at least one of the objects is still a stellar core (the other may already be a black hole). An interesting possibility that arises is that, if the orbital separation is small enough, the companion can torque the star through tides before it too becomes a black hole. This would lead to a high black hole spin aligned with the orbit. Mass transfer episodes can also spin up the objects; altogether, isolated formation generally predicts a tendency for spins to be aligned with the orbital angular momentum. Black hole spin orientations are thus of paramount importance as a discriminator between dynamical and isolated formation channels.

1.2 Thesis overview

In order to use gravitational wave data to constrain the properties of the astrophysical population, several layers of data analysis need to take place. Namely, identifying signals present in the noisy data and establishing their statistical significance, estimating their source parameters, and using the aggregated set of detections to test and inform models of the population.

1.2.1 Search pipeline

The current policy of the LIGO–Virgo Collaboration (LVC) regarding accessibility of their data is as follows. Once data are acquired, there is a proprietary period of 18 months during which these are calibrated and analyzed by the LVC. An initial search is performed at low-latency and limited information about candidate triggers is reported within minutes to hours, in order to aid follow-up of counterparts by the community of astronomers. This information includes source location and estimated probability of electromagnetic emission (e.g. by a kilonova, the explosion following a binary neutron star merger). The more refined results of an offline search are reported within the proprietary period in the form of a catalog of events, along with short segments of data that contain the events identified. After the proprietary period ends, the full strain data are released to the public (Abbott et al., 2021).

The availability of these data motivated my collaborators and me to develop a pipeline to search for compact-object binary mergers (Venumadhav et al., 2019). The task of the search pipeline is to define a detection statistic, that tests the hypothesis H_1 that there is a signal present in a piece of data against the null hypothesis H_0 that the data only contains noise. By Neyman–Pearson lemma, the optimal statistic is (a monotonic function of) the likelihood ratio

$$\frac{P(d \mid H_1)}{P(d \mid H_0)}. \tag{1.7}$$

The pipeline needs to identify all data segments where the statistic Eq. (1.7) has a high value.

A useful model to orient the discussion is that detector noise is stationary and Gaussian—this is only approximately true for real data, and can be thought as a theoretical limit for detection. Under this model, the noise likelihood

$$P(d \mid H_0) \propto \exp\left\{-\frac{1}{2}\langle d \mid d\rangle\right\}, \tag{1.8}$$

where

$$\langle a \mid b \rangle = 4\,\mathrm{Re} \int_0^\infty \mathrm{d}f \frac{\tilde{a}^*(f)\tilde{b}(f)}{S(f)} \tag{1.9}$$

is the inverse-variance-weighted inner product. Here $S(f)$ is the one-sided power spectral density (PSD) of detector noise, which is diagonal in frequency space by the stationarity hypothesis, tildes indicate Fourier transforms and asterisks complex conjugation. The signal likelihood is

$$P(d \mid H_1) = \int \mathrm{d}\theta P(d \mid \theta, H_1) P(\theta \mid H_1), \tag{1.10}$$

where θ are parameters of the gravitational wave signal and, again for stationary Gaussian noise,

$$P(d \mid \theta, H_1) \propto \exp\left\{-\frac{1}{2}\langle d - h(\theta) \mid d - h(\theta)\rangle\right\}, \tag{1.11}$$

where $h(\theta)$ is the model gravitational wave strain. For Gaussian noise, the log-likelihood ratio of a specific signal model θ is thus

$$\begin{aligned} \log \frac{P(d \mid \theta, H_1)}{P(d \mid H_0)} &= -\frac{1}{2}\langle d - h \mid d - h\rangle + \frac{1}{2}\langle d \mid d\rangle \\ &= \langle d \mid h\rangle - \frac{1}{2}\langle h \mid h\rangle. \end{aligned} \tag{1.12}$$

We can readily estimate the number of independent models θ that the pipeline needs to test. These source parameters can be divided into intrinsic (masses and spins) and extrinsic (location, orientation and time of arrival). Intrinsic parameters control the shape of the waveform, while (in the approximation that gravitational radiation is quadrupolar and that the spins are aligned with the orbit) extrinsic parameters change the amplitude, phase and time of arrival seen at each detector k:

$$\tilde{h}_k(f; \theta_{\mathrm{int}}, \theta_{\mathrm{ext}}) = A_k(\theta_{\mathrm{ext}}) e^{i\phi_k(\theta_{\mathrm{ext}})} e^{-i2\pi f t_k(\theta_{\mathrm{ext}})}\, h_0(f; \theta_{\mathrm{int}}) \tag{1.13}$$

As we will see in Chapter 2, there are $\sim 10^4$ binary black hole templates with sufficiently different shapes to be regarded as independent trials, and an order of magnitude more for the lighter binary neutron stars or neutron-star–black-hole binaries, whose signals are much longer; these cover the variability from intrinsic parameters. Typical autocorrelation times of these templates (whitened with the detector noise spectrum) are on the order of a millisecond, so there are $\sim 10^{10}$ independent possible arrival times for each template in a few-month long observing run. Depending on the sky location of the source, there is a time delay between detectors up to the gravitational wave travel

time—of $\pm 10\,\mathrm{ms}$ for Livingston–Hanford, the most sensitive pair of detectors—which multiplies the options by 20. The remaining extrinsic parameters (distance and orientation), which determine the amplitude and phase, can be analytically maximized over; for example for the amplitude

$$\max_A \log \frac{P(d \mid \theta, H_1)}{P(d \mid H_0)} = A_* \langle d \mid h_0 \rangle - \frac{1}{2} A_*^2 \langle h_0 \mid h_0 \rangle \tag{1.14}$$

$$= -\frac{1}{2} \frac{\langle d \mid h_0 \rangle^2}{\langle h_0 \mid h_0 \rangle} := -\frac{\rho^2}{2}; \tag{1.15}$$

$$A_* = \frac{\langle d \mid h_0 \rangle}{\langle h_0 \mid h_0 \rangle}; \tag{1.16}$$

$$\rho = \frac{\langle d \mid h_0 \rangle}{\sqrt{\langle h_0 \mid h_0 \rangle}}. \tag{1.17}$$

We have introduced the signal-to-noise ratio ρ, which is the cross-correlation of a normalized template $h_0/\sqrt{\langle h_0 \mid h_0 \rangle}$ with the data. For the phase, maximization is achieved by changing the real part in Eq. (1.9) by an absolute value. Maximizing over amplitude and phase means that the detection statistic follows a chi-squared distribution with two degrees of freedom on Gaussian noise.

The need for trying $10^4 \cdot 10^{10} \cdot 20 \sim 10^{15}$ models has two implications: First, it sets a limit for detectability of a signal $\rho^2_{\mathrm{min}} \approx 70$, set by

$$N_{\mathrm{trials}} \int_{\rho^2_{\mathrm{min}}}^{\infty} \mathrm{d}\rho^2 \chi^2(\rho^2, 2) \approx 1 \tag{1.18}$$

(roughly, $e^{\rho^2_{\mathrm{min}}/2} \approx N_{\mathrm{trials}}$). An optimal pipeline will be limited by the fact that, given the number of models tried, Gaussian noise will produce $\gtrsim 8\sigma$ fluctuations over the course of the observing run, this is sometimes termed look-elsewhere effect. For a sub-optimal pipeline, the detection bar will in general be higher.

Second, the sheer size of parameter space means that the cost of computing the detection statistic needs to be optimized. The solution is to introduce several stages in the pipeline which compute increasingly better approximations to the detection statistic and filter out data that, at that stage, are already inconsistent with having a detectable signal.

A first approximation consists of replacing the integral in Eq. (1.10) by a maximization over model parameters, $P(d \mid H_1) \overset{\propto}{\sim} P(d \mid \theta_*)$. Extrinsic parameters are optimized by maximizing the likelihood over amplitude, phase and time of arrival. For intrinsic parameters, we construct a bank of template waveforms that cover the target space of signals sufficiently densely and test all of them by brute force. This template bank is described in Chapter 2; its most remarkable property is that it has a geometric structure. This means that template waveforms live on a coordinate space

whose Euclidean distance corresponds to their "mismatch"—a measure of how the response of the template to a signal degrades if their shapes are different. This feature provides a few analytical and algorithmic advantages: First, a set of templates with optimal coverage can be obtained by a regular tiling of the target geometric space. Second, the search can be refined on demand around promising triggers by adding more templates locally, which decreases the overall computational cost. Finally, signal quality tests can be made orthogonal (insensitive) to mismatches that are due to the discreteness of the bank. This allows to perform more stringent tests at fixed false-rejection rate, which improves the sensitivity of the search. Another improvement we introduced with this bank is that we divide the parameter space into disjoint regions (by mass of the binary), in which we expect comparable numbers of astrophysical signals but require largely different number of templates to cover. This helps insulate the different look-elsewhere penalties associated to these regions of parameter space, as well as the different classes of background noise that predominate, and prevent these problems from cross-contaminating the rest of the search.

As discussed above, the main ingredient of the detection statistic is the cross correlation between the template and the data: whenever the data contain a signal similar to the template, the cross correlation is high. This would be the optimal detection statistic for a single detector with Gaussian, stationary noise. Since there are multiple detectors, and the noise is only approximately stationary and Gaussian, the detection statistic requires several corrections that are implemented by different stages of the pipeline as follows.

Detectors exhibit non-Gaussian transients at a rate of several per hour. These so-called "glitches" are short, abrupt disturbances whose origin is often unknown. We first identify and remove as many glitches as we can by iteratively whitening the data, running several excess power tests, and excising the glitches. The excess power tests are calibrated to guarantee a fixed, small false rejection rate of astrophysical signals with a moderate signal to noise ratio $\rho < 20$ on Gaussian noise (we assume that any louder signals have already been found in previous searches). To excise the glitches, the approach of simply gating the data with a window function would hurt the sensitivity of the search. This is because, due to narrowband lines in the detector noise power spectrum, the whitening filter has a long impulse response time of several tens of seconds. Gating would thus pollute model scores tens of seconds away from the glitch. Instead, we introduce the following change to the detection statistic (Zackay et al., 2019c). Recall from Eq. (1.17) that the signal-to-noise ratio in stationary Gaussian noise is

$$\rho = \frac{\langle d \mid h \rangle}{\sqrt{\langle h \mid h \rangle}} = \frac{h^T C^{-1} d}{\sqrt{h^T C^{-1} h}}, \tag{1.19}$$

where in the time domain the inverse of the covariance C^{-1} is a non-diagonal matrix. Our model for glitches is that they completely destroy information over a short segment in the time domain, but leave data outside it unaffected. Let us write

$$d = \begin{pmatrix} x \\ y \end{pmatrix} \tag{1.20}$$

where x denotes data that is clean, and y data that was corrupted by a glitch and became unusable; in the time domain these are well separated. The optimal detection statistic in this case becomes

$$\log \frac{P(d \mid H_1)}{P(d \mid H_0)} = \log \frac{P(x \mid H_1)}{P(x \mid H_0)} + \log \frac{P(y \mid x, H_1)}{P(y \mid x, H_0)} \tag{1.21}$$

where the last term vanishes, since our model is that glitches are allowed to do anything. For the first term we still assume that x is described by Gaussian noise under H_0, or Gaussian noise plus a signal under H_1:

$$\log \frac{P(x \mid \theta, H_1)}{P(x \mid H_0)} = -\frac{1}{2}\big(x - h_x(\theta)\big)^T C_{xx}^{-1}\big(x - h_x(\theta)\big) + \frac{1}{2} x^T C_{xx}^{-1} x \tag{1.22}$$

where $h_x(\theta)$ is the part of the template that lands on clean data and C_{xx} is the noise covariance outside the glitch. This can be mapped to the previous formulas for the detection statistic by replacing the metric C^{-1} by

$$H = \begin{pmatrix} C_{xx}^{-1} & 0 \\ 0 & 0 \end{pmatrix}; \tag{1.23}$$

for example, Eq. (1.17) becomes

$$\rho = \frac{h^T H d}{\sqrt{h^T H h}}. \tag{1.24}$$

With this technique in place, removing glitches becomes harmless and we can do it aggressively to obtain cleaner background distributions.

A second problem of real detector noise is that it is not stationary: the PSD undergoes $\sim 10\%$ variations on scales $\gtrsim 10\,\mathrm{s}$. However, due to the aforementioned narrow spectral lines, thousands of seconds of data are required in order to resolve the lines while keeping $\gtrsim 100$ independent PSD samples to keep statistical measurement errors low. Our workaround is to use a fiducial spectrum measured over few thousand seconds and then normalize the score (1.17) by its local standard deviation, empirically measured over 10^4 autocorrelation times ($\sim 10\,\mathrm{s}$) to keep statistical errors under 1% (Zackay et al., 2019c). The effect of this "PSD-drift correction" on the distribution of

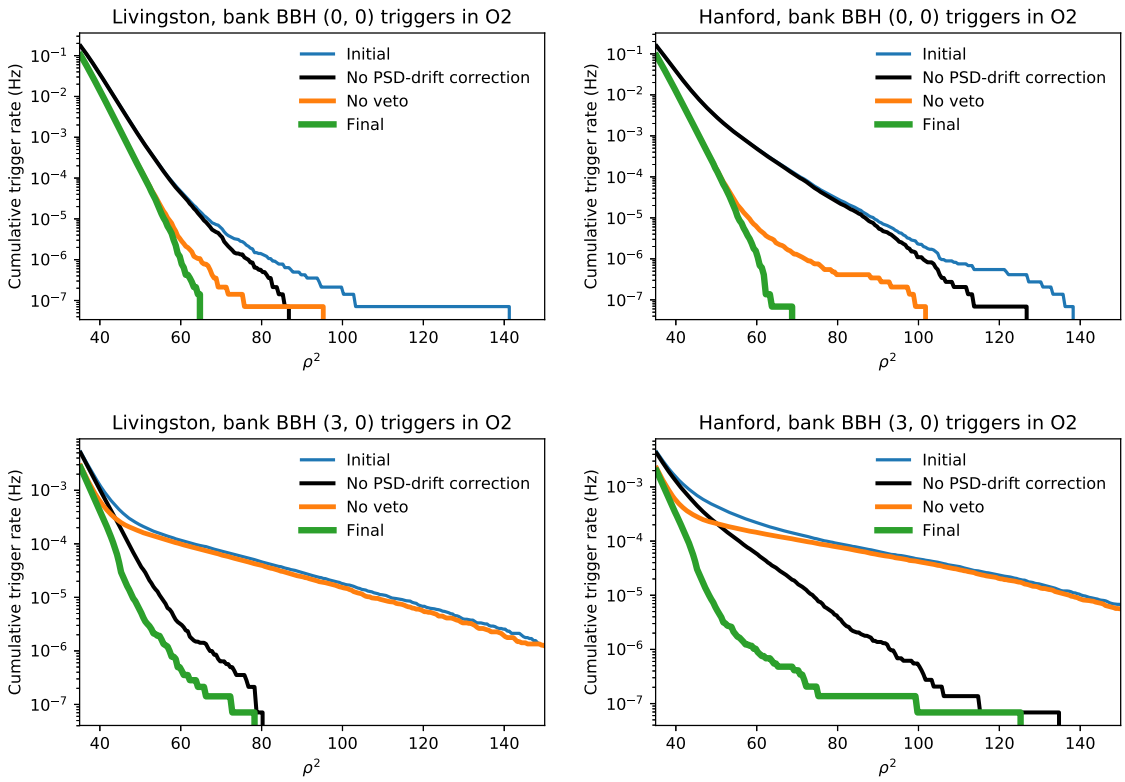


Figure 1.1: Complementary cumulative distribution of single-detector signal-to-noise ratios, for a low-mass template bank `BBH (0, 0)` (top) or a high-mass bank `BBH (3, 0)` (bottom), in the Livingston (left) or Hanford (right) detectors. Four curves are shown, toggling PSD-drift correction and vetoes on and off. These distributions are completely dominated by background noise (detected events have been removed); distributions with more support at high values of ρ^2 thus correspond to smaller sensitivity. Figure taken from Zackay et al. (2019c).

background triggers is illustrated in Fig. 1.1: it takes the blue distribution to the orange one, which more closely resembles a χ^2 distribution (a straight line in this plot). The main improvement happens at low signal to-noise-ratios.

However, especially for high-mass templates that are a few cycles long (bottom panels in Fig. (1.1)) the orange distribution still exhibits a non-Gaussian tail that would hurt sensitivity at high signal-to-noise ratios. This tail is caused by glitches that survived the generic excess power tests performed at the initial data processing stage. At this point, a series of signal-quality tests are applied to loud triggers in the form of vetoes. Since we do not have a good model for glitches, we design the tests to have a fixed false-rejection rate of $\approx 1\%$ for signals on Gaussian noise. We test that the template plus Gaussian noise is a good description of the data by checking that different parts of the template accumulate signal-to-noise ratio with the expected amplitude and phase. The template to veto is chosen by maximizing a joint score $\rho_{\rm H}^2 + \rho_{\rm L}^2$ over the two LIGO detectors, Hanford and

Livingston. This step requires refining the template bank grid around the trigger, because subtracting an incorrect template would leave non-Gaussian residuals, that could spuriously trigger the test. In addition, the tests are orthogonalized with respect to departures from the template along directions spanned by the template bank. The geometric structure of our template bank is crucial for the implementation of both these steps. The effect of vetoes on the background distribution is also shown in Fig. 1.1 by the black and green lines. For low-mass templates the distribution after PSD-drift correction and vetoes is Gaussian all the way up to 8σ. For higher masses and especially in Hanford, there remains a non-Gaussian tail that dominates above 7σ; still, the background level gets reduced by several orders of magnitude.

The next stage of the pipeline is to combine triggers from Hanford and Livingston (Virgo joined late in O2 and has a lower sensitivity than the LIGO detectors, so for simplicity we do not use it for detection). The final detection statistic is the joint log likelihood-ratio for the data at the two detectors

$$\log \frac{P(d_{\rm H}, d_{\rm L} \mid H_1)}{P(d_{\rm H}, d_{\rm L} \mid H_0)} = \log P(d_{\rm H}, d_{\rm L} \mid H_1) - \log P(d_{\rm H} \mid H_0) - \log P(d_{\rm L} \mid H_0), \tag{1.25}$$

under the assumption that $d_{\rm H}$ and $d_{\rm L}$ are uncorrelated in the noise hypothesis. The detection statistic involves a coherent score $\log P(d_{\rm H}, d_{\rm L} \mid H_1)$ and an incoherent score $-\log P(d_{\rm H} \mid H_0) - \log P(d_{\rm L} \mid H_0)$. Due to the non-Gaussian tails, we need to estimate the incoherent score from the empirical background distribution at each detector. For triggers that have a chance of reaching the detection bar, we compute the coherent score Eq. (1.10) by maximization over intrinsic parameters and Monte Carlo integration over extrinsic parameters.

Once the top triggers have been determined, there remains to estimate their significance. We do this by a bootstrap method, borrowed from LVC's PyCBC pipeline: we generate data realizations free of coincident astrophysical signals by artificially time-shifting the data between detectors beyond the 10 ms gravitational wave travel time. We then rerun the pipeline on these data and empirically reconstruct the background distribution, which we use to establish the frequentist's false-alarm rate of a trigger as a function of the value of its detection statistic.

There is one regime in which the strategy described above fails, which is when a signal is much louder in one detector than the other. This can happen for a combination of reasons: the detectors have different intrinsic sensitivities, orientations and noise realizations. The different orientations mean that for fortuitous source polarization and sky location, a detector may be more sensitive than the other to a particular source. The above strategy fails in this regime because a trigger may not

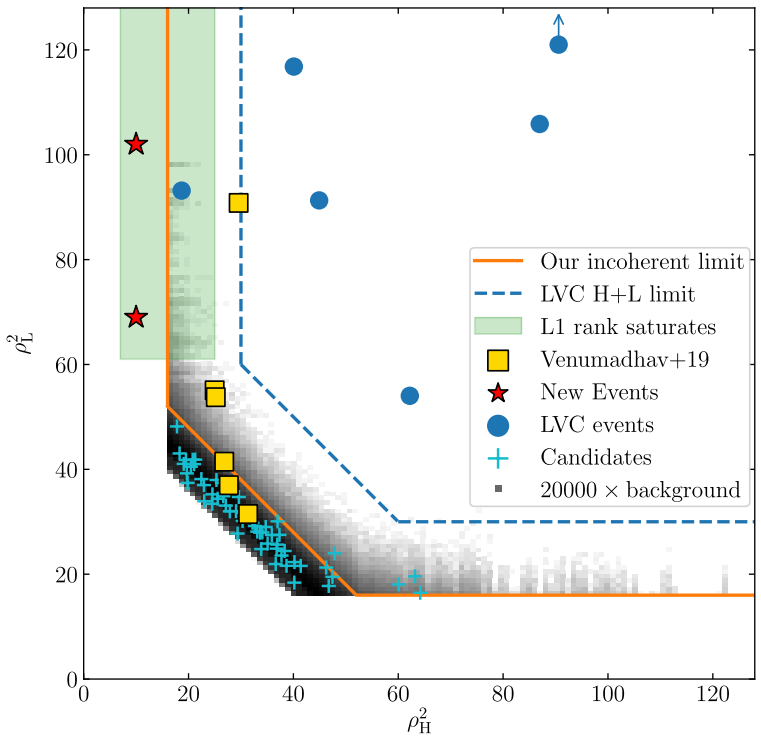


Figure 1.2: Approximate detection limits in the space of incoherent Hanford and Livingston squared signal-to-noise ratio, for high-mass ($\mathcal{M} > 20\,\mathrm{M}_\odot$) binary black holes. The detection limit of our main (2-detector coincidence) search on the O2 observing run is shown in orange, that of LVC's pipelines in blue. The green region indicates the regime in which the coincidence search may miss detectable events that are loud in Livingston and faint in Hanford, there we run a dedicated search for signals with disparate detector responses. Blue circles show events first reported by the LVC, yellow squares events first found in our coincidence search and red stars in the search for signals with disparate detector responses. A logarithmic histogram of 2×10^4 background realizations is shown in gray scale.

be found in coincidence—we impose a minimum $\rho^2 > 16$ to record a single-detector trigger, in order to keep memory usage under control. In Zackay et al. (2019a) we carried a dedicated search to cover this case, where we first ranked single-detector Livingston triggers and then directly computed their coherent score including Hanford.

1.2.2 Detections

Owing to the improvements described above, we were able to improve the sensitivity over previous searches and approximately double the sensitive volume. Figure 1.2 shows the improvement in detection limit. Our searches confirmed previous detections reported by the LVC and identified nine new signals in O1 and O2 (Zackay et al., 2019b; Venumadhav et al., 2020; Zackay et al., 2019a). We are currently analyzing the recently released first half of the third observing run, O3a.

We identified seven of these signals in Hanford–Livingston coincidence: one in O1 (GW151216,

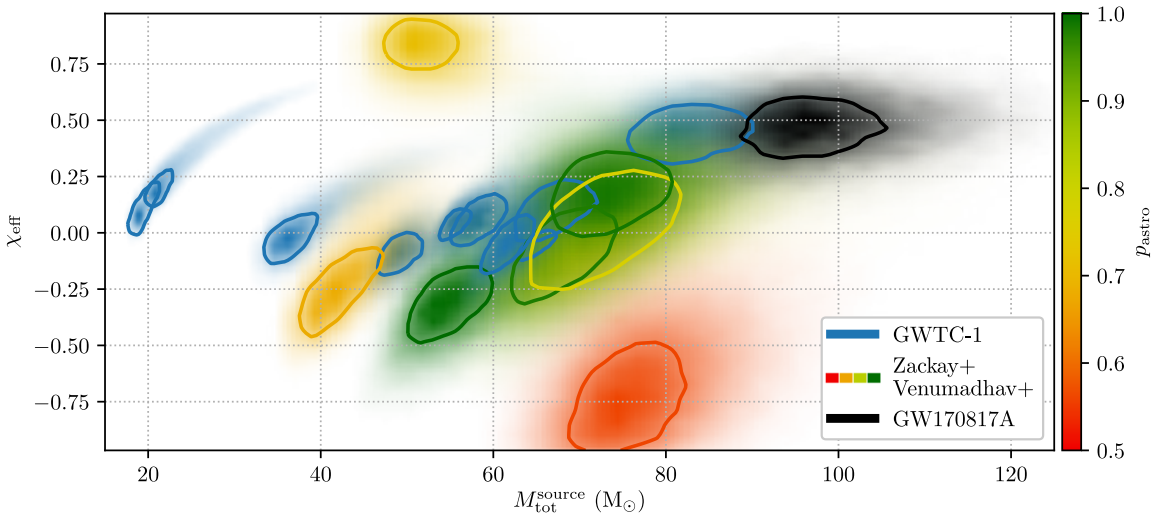


Figure 1.3: Binary black holes in O1 and O2, in terms of total mass and effective spin. Events found by the LVC are shown in blue, and those found in our coincidence search are color coded by probability of astrophysical origin. Figure from Zackay et al. (2019a).

after its birthday) and six in O2 (GW170121, GW170202, GW170304, GW170403, GW170425, GW170727). The remaining two (GW170817A and GWC170402) were found in a dedicated search for signals with disparate responses at Livingston and Hanford. Of these, the most significant are GW170727, GW170121 and GW170304, which are confidently astrophysical in origin with false-alarm rates of one per 256, 185 and 79 O2 observing runs, respectively. The other events have more modest inverse false-alarm rates, which means it is possible that some of them are due to detector noise. Events from the first two observing runs are shown in Fig. 1.3 in terms of their total mass $M_{\rm tot} = m_1 + m_2$ and effective spin

$$\chi_{\rm eff} = \frac{m_1 \boldsymbol{\chi_1} + m_2 \boldsymbol{\chi_2}}{m_1 + m_2} \cdot \hat{\boldsymbol{L}}, \tag{1.26}$$

where $\boldsymbol{\chi} = Gm^2\boldsymbol{S}/c$ is the dimensionless spin and $\hat{\boldsymbol{L}}$ is the direction of the orbital angular momentum. Events found in our coincidence search are color-coded by their probability of astrophysical origin.

In terms of their parameters, notable events include GW170121, which shows a preference for spin anti-aligned with the orbit ($\chi_{\rm eff} < 0$) at 95% credibility. GW170304 also has anti-aligned spins, but the event itself has a smaller significance. GW151216, conversely, requires large spins with a high degree of alignment with the orbit. This is interesting because it is very unlikely to get both spins aligned from the isotropic distribution expected in dynamical channels, while other formation channels do have mechanisms for producing large aligned spins (e.g. through tides). Another remarkable event is GW170817A, which at the time was the most massive event detected

and also has a moderately large spin aligned with the orbit. In the O3a run, more events similar to it were found, which increases our confidence in its astrophysical origin. Finally, GWC170402 is interesting in that, despite its rather low false-alarm rate, we were unable to find a physical model that describes well the data. We found two different physical solutions with comparable signal-to-noise ratio, which means that an unphysical linear combination of these fits the data significantly better. This might be due to waveform systematics like physical effects missing from our waveform models (e.g., eccentricity), or simply to a rare non-Gaussian noise transient.

1.2.3 Population inference

Once signals have been identified, their individual source parameters need to be estimated. Often there are large measurement uncertainties and correlations in these parameters; in addition, gravitational wave detectors have nontrivial selection effects (for example, from Eq. (1.6) we already see that heavier systems emit louder gravitational waves and can be observed to further distances). These complications need to be accounted for in order to get unbiased constraints. Chapters 3, 4 and 5 describe a series of works on this area, carried out as events from O2 were being announced by the LVC (2018–2019), after our O1 and O2 reanalyses (2019–2020), and after the LVC announced events from O3a (2020–2021), respectively. Each iteration progressively involved more events, required more sophisticated methods and yielded more insights.

The analysis in Chapter 3 is based on events reported in the first Gravitational Wave Transient Catalog by the LVC (GWTC-1, Abbott et al. (2019a)), which contains 10 binary black hole events from O1 and O2. At that time, short segments of data around each event had been released. I introduce some approximations that help develop intuition and allow to marginalize the likelihood function of individual event parameters semi-analytically. I then develop a method to combine these into a likelihood function for population parameters, which I use to place constraints on the distribution of astrophysical binary black hole spins and masses.

In Chapter 4, I carry a full end-to-end population analysis of the results of our search for binary black holes in O1 and O2. Some events in our catalog have marginal significance, in the sense that there is a non-negligible probability that they are due to detector noise instead of astrophysical signals. I develop a novel framework that enables inclusion of candidate events with arbitrary significance in the population inference. This allows to extract the information that these marginal triggers collectively carry about the astrophysical population without introducing biases. Applying this framework to our catalog, I find improved constraints on the spin, mass, rate and cosmological

evolution of the population of binary black holes.

In Chapter 5, I incorporate the GWTC-2 catalog, which includes events from O3a identified by the LVC, in the population inference. Due to improvements in detector sensitivity, the number of events tripled with the inclusion of GWTC-2. With this larger statistical sample, tests for qualitative, robust features of the distribution of black hole spin orientations predicted by different classes of formation channels become powerful enough to place meaningful bounds. Most interestingly, I find that there is an excess of black hole spins aligned with the orbit over anti-aligned, which disfavors dynamical formation as an explanation for the entirety of the events.

Chapter 2

Template Bank

This Chapter has been adapted from Roulet et al. (2019).

We introduce an algorithm for placing template waveforms for the search of compact binary mergers in gravitational wave interferometer data. We exploit the smooth dependence of the amplitude and unwrapped phase of the frequency-domain waveform on the parameters of the binary. We group waveforms with similar amplitude profiles and perform a singular value decomposition of the phase profiles to obtain an orthonormal basis for the phase functions. The leading basis functions span a lower-dimensional linear space in which the unwrapped phase of any physical waveform is well approximated. The optimal template placement is given by a regular grid in the space of linear coefficients. The algorithm is applicable to any frequency-domain waveform model and detector sensitivity curve. It is computationally efficient and requires little tuning. Applying this method, we construct a set of template banks suitable for the search of aligned-spin binary neutron star, neutron-star–black-hole and binary black hole mergers in LIGO–Virgo data.

2.1 Introduction

The optimal algorithm to search for known signals in the presence of Gaussian noise is matched-filtering, in which a signal template is cross-correlated with the data and triggers are recorded whenever the correlation exceeds some threshold. In the context of gravitational wave detection with the LIGO (Aasi et al., 2015) and Virgo (Acernese et al., 2014) interferometers, compact binary coalescences are a good example of predictable signals for which we have accurate models, and thus are well suited for matched filtering (Dhurandhar and Sathyaprakash, 1994; Allen et al., 2012). Indeed, the LIGO and Virgo Collaborations have reported gravitational wave signals from 10 binary

black hole (BBH) and one binary neutron star (BNS) mergers during their first and second observing runs (Abbott et al., 2019a, 2016a,b,c, 2017b,c,d,e), all of which were found by search pipelines based on matched-filtering (Sachdev et al., 2019; Usman et al., 2016) (seven of these BBHs were also found by an unmodeled search (Klimenko et al., 2016; Abbott et al., 2019a)). Searches in the public LIGO–Virgo data by independent groups have found seven additional BBHs (Zackay et al., 2019b; Venumadhav et al., 2020), additional BBH candidates (Nitz et al., 2019a) and a BNS candidate (Nitz et al., 2019b), also employing matched filtering.

Since the source parameters describing the waveform are not known *a priori*, one needs a bank of waveform templates that adequately cover the parameter space. The notion of good coverage is characterized by the warranty that any physical waveform the search aims to detect has a sufficiently large match with at least one waveform in the bank. For example, the LIGO and Virgo Collaborations have aimed at a minimum match of 97% for any aligned-spin binary merger with component masses between 1 and $\sim 200\,\mathrm{M}_\odot$, for which they require a template bank consisting of $\sim 4 \times 10^5$ waveforms (Canton and Harry, 2017). Due to the large number of templates involved, matched filtering is a sizeable computational task. This means that an efficient bank should not over-cover the parameter space. In other words, the templates should be uniformly spaced with respect to a distance defined in terms of the matched-filtering mismatch between templates (defined in §2.2). This incorporates the notion that, from the perspective of signal detection, two waveforms that sufficiently resemble each other are essentially indistinguishable in the presence of noise. Source parameters can be mutually degenerate in the sense that different parameter combinations may describe similar waveforms. The optimal placement of templates in physical parameter space is very non-uniform; for example, an order of magnitude more templates are needed to search for mergers with 1–3 $\mathrm{M}_\odot$ components ("neutron stars") than for mergers with 3–200 $\mathrm{M}_\odot$ ("black holes").

Two broad classes of template-placement algorithms have been developed in the literature. One robust method is "stochastic placement" (Harry et al., 2009; Ajith et al., 2014; Privitera et al., 2014; Capano et al., 2016): waveforms are randomly drawn from the desired parameter space, and one gradually builds up the bank by only accepting newly drawn waveforms that differ sufficiently from the ones the bank already has, and rejecting those that are too similar to at least one existing waveform. Stochastic placement, however, has the shortcoming that a large number of trial waveforms needs to be drawn before convergence is achieved (much more than the required number of templates in the bank). This method also tends to over-cover the parameter space, in the sense that the average template density is higher than optimal at fixed minimum match (Roy et al., 2017).

A different method to construct the bank is "geometric placement". Here, a metric in the

parameter space is defined based on the matched-filtering overlap between waveforms (Owen, 1996; Owen and Sathyaprakash, 1999). This metric is then used to define a regular lattice (Babak et al., 2006; Cokelaer, 2007; Babak et al., 2013). However, it is in general difficult to derive this metric, especially if the parameter space is high dimensional or if the waveform model is not analytic. Approximations to the metric have first been found by using suitably reparameterized analytic, post-Newtonian (PN) non-spinning waveform models (Owen, 1996; Owen and Sathyaprakash, 1999; Tanaka and Tagoshi, 2000); later generalizations include the use of phenomenological waveform models and template parameters (Ajith et al., 2008), the inclusion of aligned-spin PN models (Brown et al., 2012; Harry et al., 2014), or numerical evaluation from arbitrary waveform models (Roy et al., 2019).

In practice, a combination of the two methods is often a better strategy. For example, one can place templates geometrically at low masses and stochastically at high masses (Capano et al., 2016; Canton and Harry, 2017), or one can use many small patches with regularly spaced templates, which are themselves placed stochastically to cover the entire parameter space (Roy et al., 2019).

In this work, we develop a fast and general method to construct a high-effectualness template bank using geometric placement. Our method relies on the construction of a flat, linear space of orthonormal phase functions that embeds the space of physical waveforms. The Euclidean distance in this space coincides with the mismatch distance between similar waveforms, making these coordinates naturally suited for geometric placement of templates. Besides optimal template placement, having this geometrical notion turns out to be helpful for a number of reasons. It allows to refine the bank locally around triggers at the time of search, reducing the amount of templates in the bank at fixed effective coverage. Moreover, a crucial stage of searches involves signal consistency checks, that assess the probability that the residual between a best-fitting template and a candidate signal is explained by Gaussian noise in order to reject non-Gaussian noise transients (Allen, 2005; Sachdev et al., 2019; Usman et al., 2016; Venumadhav et al., 2019). With the bank described here, these tests can be made orthogonal to the linear space of waveforms, so that they are insensitive to mismatches due to the discreteness of the bank. This allows to make the tests more stringent and improves the sensitivity of the search (Venumadhav et al., 2019). We further require that the template bank be built from sub-banks that can be approximated to have a fixed amplitude profile $A(f)$. This feature is useful for implementing the noise amplitude-spectral-density drift correction, a key component for precise matched filtering (Venumadhav et al., 2019; Zackay et al., 2019c). Together, these analytical properties make our template bank appealing, even considering that there are other template banks with comparable effectualness and number of templates in the literature. Finally, building a new

template bank enables us to customize a number of other choices, like the frequency range and parameter space covered, in the context of our search pipeline (Venumadhav et al., 2019) and the detector performances during the observation time analyzed. The coordinates presented in this work are similar in essence to the ones introduced in Brown et al. (2012), except that we generalize them to arbitrary waveform models and component mass ranges.

The paper is organized as follows. In §2.2 we define a metric based on the mismatch between templates and show how the desired Euclidean space can be constructed. In §2.3 we apply this formalism to the construction of a search-quality template bank that targets stellar-mass compact binary mergers. We summarize our results in §2.4. The bank presented here was used in the searches described in Refs. (Venumadhav et al., 2019, 2020).

2.2 Linear metric space

In this section, we define the notion of distance between templates and describe the construction of a low-dimensional linear space of phase functions in which the metric is Euclidean. We build this linear space based on the intuition that the unwrapped phases are smooth functions of the wave frequency (Cutler and Flanagan, 1994) and hence are linear combinations of a small number of basis functions (Tanaka and Tagoshi, 2000; Brown et al., 2012).

2.2.1 Mismatch distance

We first introduce the noise-weighted inner product in the frequency domain (Allen et al., 2012)

$$(h_i \mid h_j) := 4 \int_0^\infty \mathrm{d}f \, \frac{\tilde{h}_i(f)\tilde{h}_j^*(f)}{S_n(f)}. \tag{2.1}$$

Here, $S_n(f)$ is a fiducial one-sided power spectral density (PSD) of the detector noise and tildes indicate Fourier transforms. The match between h_i and h_j is given by $\mathrm{Re}\, z_{ij}$, where

$$\begin{aligned} z_{ij} &:= \frac{(h_i \mid h_j)}{\sqrt{(h_i \mid h_i)(h_j \mid h_j)}} \\ &\equiv (h_i \mid h_j). \end{aligned} \tag{2.2}$$

In the second line, we normalize the waveforms to

$$(h \mid h) = 1, \tag{2.3}$$

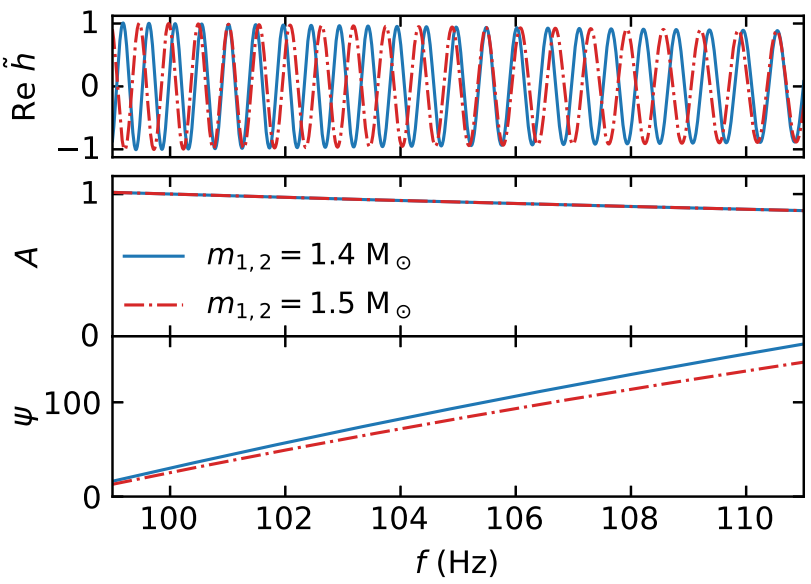


Figure 2.1: An example of two waveforms that look very different to each other in the frequency domain (top panel) but have very similar amplitude and phase profiles (middle and bottom panels). The amplitude and phase profiles can be well captured by a low-dimensional linear space spanned by a few basis functions. Waveform amplitudes are shown in arbitrary units.

as usually the template waveforms are defined up to an overall normalization. Since all possible coalescence times and phases are searched for, waveforms related by time and phase offsets are described by the same waveform in template bank. Thus, the match is maximized over time and phase offsets:

$$\begin{aligned} m_{ij} &:= \max_{\tau_0,\phi_0} \left(\mathrm{Re}\, z_{ij}(\tau_0,\phi_0) \right) \\ &= \max_{\tau_0} |z_{ij}(\tau_0)|, \end{aligned} \tag{2.4}$$

where τ_0 and ϕ_0 are the time and phase offsets between h_i and h_j, respectively. We define the mismatch distance d_{ij} between the two waveforms by

$$d_{ij}^2 = 1 - m_{ij}. \tag{2.5}$$

We seek a parametrization of waveforms under which the mismatch distance has an Euclidean metric for similar waveforms.

2.2.2 Linear space

A general frequency-domain waveform model can be cast to the form

$$\tilde{h}(f;\boldsymbol{p}) = A(f;\boldsymbol{p})e^{i\psi(f;\boldsymbol{p})}. \tag{2.6}$$

Under the approximation that the dominant mode of gravitational radiation has $(\ell, |m|) = (2, 2)$ and that spin-orbital precession and eccentricity effects are insignificant, the frequency dependent functions A and ψ vary slowly with the binary parameters $\boldsymbol{p}$, as illustrated in Fig. 2.1. For matched filtering, the phase ψ is the most important to describe with high accuracy, since loss of phase coherence leads to a rapid degradation of the signal-to-noise ratio (SNR).

Moreover, it is important to analyze templates with different amplitude profiles $A(f)$ separately as the matched filtering correction from PSD drifts depends on $A(f)$ (Zackay et al., 2019c). Thus we assume in the following that $A(f;\boldsymbol{p}) \approx \overline{A}(f)$ is valid for a suitably chosen subset of parameters. To achieve this, we sort a large number of randomly sampled physical input waveforms into groups of similar amplitude profile. In each group, we require that the match of the amplitudes to a reference $\overline{A}(f)$ exceeds a minimum

$$(A_i \mid \overline{A}) = 4\int_0^\infty \mathrm{d}f \frac{A_i(f)\overline{A}(f)}{S_n(f)} \geqslant 0.96 \tag{2.7}$$

for all input waveforms h_i in the group. Note that the match of the amplitudes sets an upper bound on the match of the waveforms. Our approach will be to split a template bank into "subbanks", each subbank describing one group of input waveforms which share the same approximate amplitude profile $\overline{A}(f)$.

We design the subbanks in order to minimize the average amplitude mismatch as follows. We start with a single subbank that contains all the input waveforms, and define its reference amplitude profile as the root-mean-square

$$\overline{A}(f) \coloneqq \sqrt{\langle A^2 \rangle(f)}, \tag{2.8}$$

where the angled brackets indicate average over the input waveforms in the subbank. This choice inherits the normalization of the input waveforms. We compute the amplitude match Eq. (2.7) for all the waveforms; if the worst match satisfies the chosen bound we stop. If it does not, we add a new subbank with a reference amplitude given by the waveform with the worst amplitude match. We then optimize the choice of reference amplitudes using the k-means algorithm: we reassign waveforms to subbanks by their best amplitude match, redefine the amplitude profile of the subbanks using Eq. (2.8), and iterate these two steps a few times to achieve convergence. Finally we recompute the

worst match and decide if a new subbank is needed, in which case we repeat the above process.

Having decided on the division of subbanks, we wish to find an efficient representation of the set of phases $\psi(f)$ as a linear combination of a small number of basis functions,

$$\psi(f;\boldsymbol{p}) = \overline{\psi}(f) + \sum_{\alpha}^{\text{few}} c_\alpha(\boldsymbol{p})\psi_\alpha(f), \tag{2.9}$$

where α is an integer index that enumerates the basis functions and $\overline{\psi}(f)$ is an average phase which we are free to define. From now on, we abandon the physical parameters $\boldsymbol{p}$ and describe the waveforms in terms of their c_α components:

$$h(f;\boldsymbol{c}) = \overline{A}(f)\exp\Big[i\Big(\overline{\psi}(f) + \sum_\alpha c_\alpha\,\psi_\alpha(f)\Big)\Big]. \tag{2.10}$$

We now express the match between two waveforms using the above decomposition. As mentioned earlier, template waveforms are defined up to arbitrary time and phase offsets, namely an additive piece to the phase that is a linear function of the frequency $\Delta\psi(f) = \phi_0 + 2\pi f\tau_0$. We choose the first two basis functions $\psi_0(f)$ and $\psi_1(f)$ to span the subspace of linear phases so that c_0 and c_1 capture phase and time offsets, respectively, and in particular $\psi_0(f) \equiv 1$. If two waveforms are similar, their inner product Eq. (2.1) to second order in δc_α is approximately

$$\begin{aligned}
&\big(h(\boldsymbol{c})\,\big|\,h(\boldsymbol{c}+\boldsymbol{\delta c})\big) \\
&\quad= 4\int_0^\infty \mathrm{d}f\,\frac{\overline{A}^2(f)}{S_n(f)}e^{i\sum_\alpha \delta c_\alpha\psi_\alpha(f)} \\
&\quad\approx 4\,e^{i\delta c_0}\int_0^\infty \mathrm{d}f\,\frac{\overline{A}^2(f)}{S_n(f)}\Big[1 + i\sum_{\alpha\geqslant 1}\delta c_\alpha\,\psi_\alpha(f) - \frac{1}{2}\sum_{\alpha,\beta\geqslant 1}\delta c_\alpha\,\delta c_\beta\,\psi_\alpha(f)\,\psi_\beta(f)\Big] + \mathcal{O}(\delta c^3).
\end{aligned} \tag{2.11}$$

This motivates a new inner product, with respect to which we will orthonormalize the basis functions:

$$\begin{aligned}
\langle\psi_\alpha,\psi_\beta\rangle &:= 4\int_0^\infty \mathrm{d}f\,\frac{\overline{A}^2(f)}{S_n(f)}\,\psi_\alpha(f)\,\psi_\beta(f) \\
&\equiv \delta_{\alpha\beta},
\end{aligned} \tag{2.12}$$

which we enforce by a suitable choice of the basis functions $\psi_\alpha(f)$ (described below). In particular,

the first condition $\langle\psi_0,\psi_0\rangle = 1$ is the normalization Eq. (2.3), and the two first basis functions are

$$\begin{aligned}\psi_0(f) &= 1,\\ \psi_1(f) &= \frac{f-\overline{f}}{\sqrt{\overline{f^2}-\overline{f}^2}},\end{aligned} \tag{2.13}$$

where we define $\overline{f^n} \coloneqq 4\int_0^\infty \mathrm{d}f f^n A^2(f)/S_n(f)$.

Using orthonormality, Eq. (2.11) becomes

$$\big(h(\boldsymbol{c})\,\big|\,h(\boldsymbol{c}+\boldsymbol{\delta c})\big) \approx e^{i\delta c_0}\Big(1-\frac{1}{2}\sum_{\alpha\geqslant 1}\delta c_\alpha^2\Big)+\mathcal{O}(\delta c^3). \tag{2.14}$$

Thus, for nearby templates the distance Eq. (2.5) is

$$d^2_{\boldsymbol{c},\boldsymbol{c}+\boldsymbol{\delta c}} \approx \frac{1}{2}\sum_{\alpha\geqslant 2}\delta c_\alpha^2+\mathcal{O}(\delta c^3), \tag{2.15}$$

which means that the mismatch distance is given by an Euclidean metric in $\boldsymbol{c}$ space at small displacements. We construct the bank on a regular grid in $\boldsymbol{c}$ space with spacings $\Delta c_\alpha \lesssim 1$, chosen sufficiently small so as to guarantee a minimal loss of match.

We note in passing that we can also compute the distance in the opposite limit of large separation, which is useful for estimating the long-range correlations between triggers from different templates during a search. Assuming now that the templates are separated by $\boldsymbol{\delta c} = D\hat{\boldsymbol{n}}$, with $\sum_\alpha n_\alpha^2 = 1$ and $D \gg 1$, we can perform a stationary phase approximation around the frequencies f_j at which $\sum_\alpha n_\alpha \psi'_\alpha(f_j) = 0$. This yields

$$\begin{aligned}\big(h(\boldsymbol{c})\,\big|\,h(\boldsymbol{c}+D\hat{\boldsymbol{n}})\big) &= 4\int_0^\infty \mathrm{d}f \frac{\overline{A}^2(f)}{S_n(f)}\exp\left(iD\sum_\alpha n_\alpha\psi_\alpha(f)\right)\\ &\approx \frac{4}{\sqrt{D}}\sum_j \frac{\overline{A}^2(f_j)}{S_n(f_j)}\frac{\sqrt{2\pi}\exp\big(-i\frac{\pi}{4}+iD\sum_\alpha n_\alpha\psi_\alpha(f_j)\big)}{\sqrt{\sum_\alpha n_\alpha\psi''_\alpha(f_j)}}.\end{aligned} \tag{2.16}$$

Thus, the long-range correlation between two templates separated by D decays as $1/\sqrt{D}$ (this holds for the match without maximization over time).

In practice we choose the set of basis functions $\psi_\alpha(f)$ as follows:

1. Define a discrete frequency grid $\{f_k\}$ (our choice is described in §2.3). The integrals over frequency will be approximated by quadratures $\sum_k \Delta f_k \ldots$;

2. Compute a moderately large number of waveforms for random parameter choices (we use

5×10^4), and extract the unwrapped phases, $\{\psi^{(i)}(f_k)\}$, as illustrated in the top panel of Fig. 2.2;

3. Subtract the average phase $\overline{\psi}(f)$;

4. Subtract the projection onto the first two dimensions so that the phase residuals

$$\delta\psi^{(i)}(f) = \psi^{(i)}(f) - \overline{\psi}(f) - \sum_{\alpha=0}^{1} \langle \psi^{(i)} - \overline{\psi}, \psi_\alpha \rangle \, \psi_\alpha(f) \tag{2.17}$$

are orthogonal to ψ_0 and $\psi_1(f)$ with respect to the inner product Eq. (2.12) (middle panel of Fig. 2.2);

5. Construct a matrix of weighted phase residuals

$$\begin{aligned} M_{ik} &= w_k \, \delta\psi^{(i)}(f_k), \\ w_k &= 2 \, \overline{A}(f_k) \, \sqrt{\Delta f_k / S_n(f_k)}, \end{aligned} \tag{2.18}$$

and find its singular-value decomposition (SVD)

$$M_{ik} = \sum_\alpha U_{i\alpha} D_\alpha V_{\alpha k}. \tag{2.19}$$

U, V are orthogonal matrices and we sort the axes so that the eigenvalues $D_\alpha > 0$ are in decreasing order. From the orthogonality of V, i.e. $\sum_k V_{\alpha k} V_{\beta k} = \delta_{\alpha\beta}$, we can identify

$$V_{\alpha k} = w_k \, \psi_\alpha(f_k) \tag{2.20}$$

which satisfies the orthonormality Eq. (2.12) and defines the basis functions, with the convention that the α start at 2 (bottom panel of Fig. 2.2).

From Eqs. (2.9) and (2.19) it follows that the components of the input waveforms are

$$c_\alpha^{(i)} = U_{i\alpha} \, D_\alpha. \tag{2.21}$$

Since U is an orthogonal matrix, $|U_{i\alpha}| \leqslant 1$ and $|c_\alpha| \leqslant D_\alpha$, that is, the extent spanned by the input samples along each dimension in component space is bounded by D_α. This means that the information in the templates is captured by the first few components along the larger dimensions, and we can reduce the dimensionality of our description by dropping dimensions that have $D_\alpha \ll 1$.

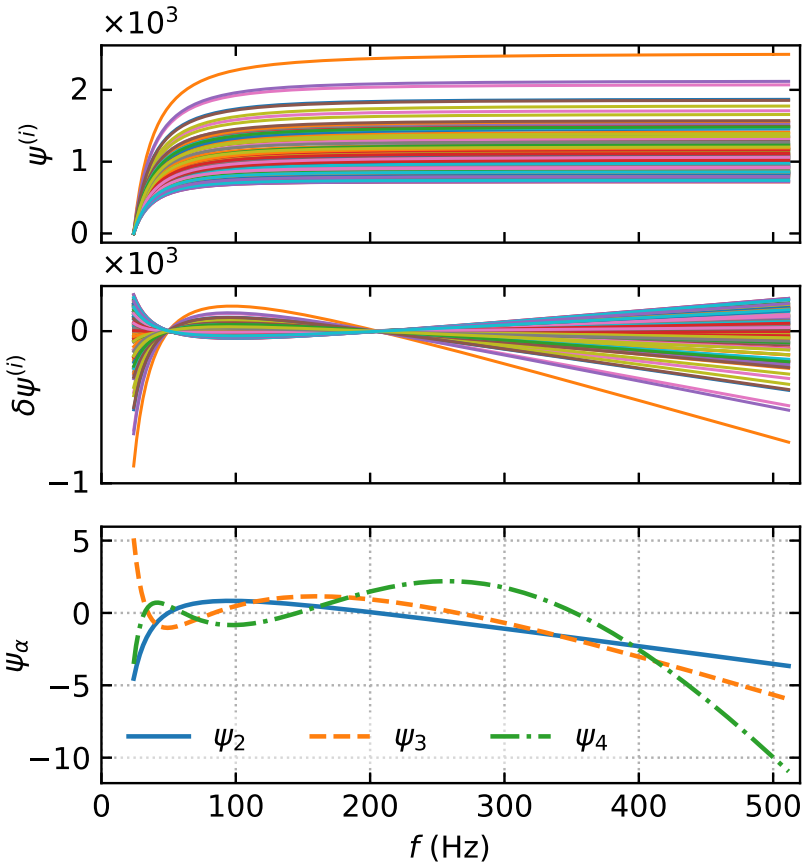


Figure 2.2: Construction of the basis functions ψ_α. *Top panel:* (a subset of 100) input unwrapped phases for random parameters. *Middle panel:* phase residuals after subtraction of the average phase and orthogonalization with respect to time and phase offsets. *Bottom panel:* first three basis functions.

2.3 Constructing a search quality template bank

In this Section, we apply the method developed in §2.2 to the construction of a template bank suitable to the search of gravitational wave strain signals from binary neutron stars, neutron-star–black-hole and binary black hole mergers.

We choose lower and upper frequency cutoffs of $f_{\rm min} = 24\,{\rm Hz}$, and $f_{\rm max} = 512\,{\rm Hz}$, respectively. These cutoffs are chosen such that the resulting loss in SNR^2 is lower than 2% for binary neutron star templates (the amplitude profiles of these whitened waveforms, i.e., $A(f)/\sqrt{S_n(f)}$, are essentially independent of parameters since the cut-off scale is outside the LIGO sensitivity band). Formally, the accumulated $\mathrm{SNR}^2 = 4\int_0^{24\,{\rm Hz}} {\rm d}f A^2(f)/S_n(f) \approx 4\int_{512\,{\rm Hz}}^{\infty} {\rm d}f A^2(f)/S_n(f) \approx 10^{-2}$ outside our frequency range. It is advisable to restrict the frequency range because the linear-free phase, and thus the basis functions, grow rapidly at both ends (see Fig. 2.2), and our Taylor expansion Eq. (2.11) would become inaccurate. As we noted above, it is exactly at these frequencies where the contribution to the matched-filtering SNR vanishes. It is better to discard these frequencies rather than to try and capture the negligible information content within by adding extra dimensions to the template bank. Furthermore, this has the additional benefit that the strain data can be down-sampled during analysis, which reduces the computational cost of the search.

We define the fiducial PSD empirically from the PSDs of 200 LIGO Handford and LIGO Livingston data files chosen randomly from the Second Advanced LIGO Observing Run (O2) release (Abbott et al., 2021). Each individual PSD was computed as described in (Venumadhav et al., 2019). The fiducial PSD is constructed using the 10th percentile of all the sample PSDs in each frequency bin. This choice is robust to large fluctuations in the sample PSDs, and is representative of optimal detector conditions.

We choose a target parameter space of compact binary mergers satisfying the following bounds:

$$1\,{\rm M}_\odot < m_2 < m_1 < 100\,{\rm M}_\odot, \tag{2.22}$$

$$q > \begin{cases} 1/50 & \text{if } m_2 < 3\,{\rm M}_\odot \\ 1/18 & \text{otherwise,} \end{cases} \tag{2.23}$$

$$|\chi_{1,2}| < 0.99, \tag{2.24}$$

where m_1 and m_2 are the primary and secondary masses, respectively, $q = m_2/m_1 \leqslant 1$ is the mass ratio, and χ_1 and χ_2 are the individual dimensionless spin projections in the direction of the orbital angular momentum. The parameter ranges and approximant used are not a constraint from the

LIGO and Virgo detectors or the method presented here, but a documentation of the choices we made. In particular, the mass ratio cut for BBHs was due to the calibration regime of the `IMRPhenomD` approximant (Khan et al., 2016). For NSBH, we extend the maximal q because a substantial part of the NSBH parameter space lies outside the calibrated range of `IMRPhenomD`. For the purpose of signal detection (as opposed to parameter estimation), the calibration tolerance is less stringent, as long as a signal can be recovered by the model with some combination of parameters. Compared to other template banks in the literature, the one presented here covers a larger spin range for low-mass objects. Indeed, bounds of $|\chi| < 0.05$ (Canton and Harry, 2017; Brown et al., 2012; Roy et al., 2019) or $|\chi| < 0.4$ (Brown et al., 2012) have been used in the BNS mass range, the former motivated by the known binary neutron star spins and the latter by the known pulsar spins (Miller and Miller, 2015). Neutron stars can in principle have dimensionless spins up to a mass-shedding limit of $|\chi| \sim 0.7$ (Lo and Lin, 2011; Tacik et al., 2015). Other types of compact objects, including light black holes, may in principle have even higher spins. This motivates us to cover this unexplored part of the parameter space.

As mentioned before, the number of templates required to describe waveforms from low-mass mergers is significantly larger than that for high-mass mergers, due to the larger number of wave cycles in band. Searches with larger template banks suffer a penalty in sensitivity because of the increased look-elsewhere effect. To prevent the high penalty inherent to the lower-mass region of parameter space from affecting the higher-mass regions, we propose to divide the search space into a number of regions and perform an independent search in each. Each search then only pays an additional look-elsewhere penalty that a few other searches are performed, but is unaffected by the potentially huge size of the other banks. This division can be interpreted as implementing a prior about which templates are more likely to produce an astrophysical trigger: if we expect comparable numbers of high- and low-mass signals but have vastly more templates at low-mass, any particular low-mass template is much less likely to produce an astrophysical trigger. In addition, templates in different regions of parameter space are sensitive to different types of noise transients in the strain data. Dividing the search into several regions enables us to recognize the different types of noise background that a search using each class of templates is subject to.

Under the above motivations, we divide the search space into regions based on the component masses, and construct a separate template bank for each of them. The division is illustrated in Fig. 2.3 and is defined as follows. We refer to binary components with masses between 1 and $3\,\mathrm{M}_\odot$ as neutron stars, and to components with masses between 3 and $100\,\mathrm{M}_\odot$ as black holes. We make three binary neutron star template banks, three neutron-star–black-hole (NSBH) banks, and five

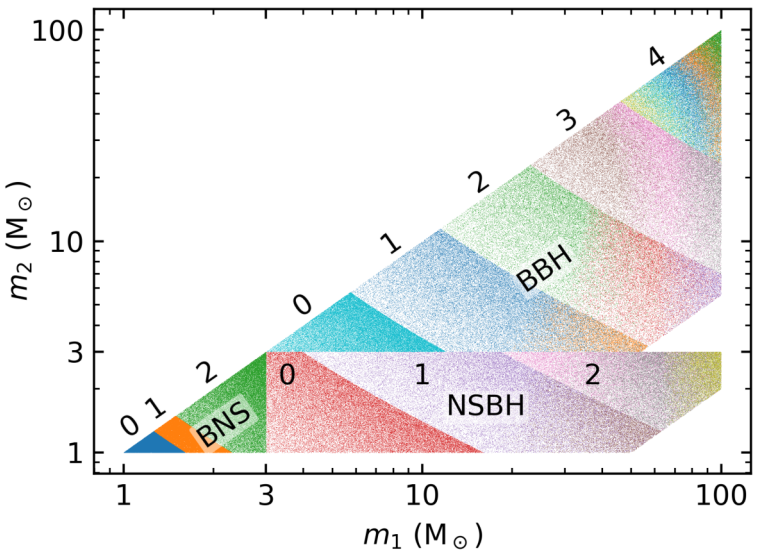


Figure 2.3: Component masses of the input waveforms used for constructing our template banks. We divide the parameter space according to the component masses into three banks of binary neutron star waveforms (`BNS 0-2`), three banks of neutron-star–black-hole waveforms (`NSBH 0-2`), and five banks of binary black hole waveforms (`BBH 0-4`). We further divide each bank into subbanks (color coded) under the criterion that the match of the amplitude profiles with a reference profile specific to each subbank exceeds 0.96. The dots show the 5×10^4 input waveforms we use to build each bank. We do not scatter plot here individual template waveforms in each bank, as those lack an explicit representation in terms of the source physical parameters.

binary black hole banks. The banks within each of these categories are defined by bins in the chirp mass $\mathcal{M} \coloneqq (m_1\, m_2)^{3/5}/(m_1 + m_2)^{1/5}$. We put the bounds between the three BNS banks at $\mathcal{M} = \{1.1, 1.3\}\,\mathrm{M}_\odot$. This choice is motivated by the observation that the chirp masses of the known Galactic binary neutron stars expected to merge within a Hubble time lie in a narrow range (Farrow et al., 2019), and therefore we might expect more astrophysical signals from this chirp mass range (which we further expand to account for the redshift of the detector-frame masses up to $z \sim 0.05$, or a luminosity distance $d_L \sim 200\,\mathrm{Mpc}$). In this way, we minimize the number of templates in the most astrophysically probable BNS bank, `BNS 1`, enhancing our sensitivity to those systems. A similar strategy was adopted in Refs. (Magee et al., 2019; Nitz et al., 2019b). For other banks, we use logarithmic chirp-mass bins: we place the bounds between the three NSBH banks at $\mathcal{M} = \{3, 6\}\,\mathrm{M}_\odot$, and those between the five BBH banks at $\mathcal{M} = \{5, 10, 20, 40\}\,\mathrm{M}_\odot$. We generate 5×10^4 input waveforms in each bank using the `IMRPhenomD` approximant (Khan et al., 2016). Based on the amplitude profiles $A_i(f)$ of the input waveforms, we further divide each bank into subbanks as explained in §2.2. We find that a single subbank is sufficient for waveforms with $m_{1,2} \lesssim 15\,\mathrm{M}_\odot$, but multiple amplitude subbanks are needed for heavier mergers as the frequency at which $A(f)$ is cut-off falls within the LIGO sensitive band. Table 2.1 summarizes the parameters of all template

Bank	$m_1\,(\mathrm{M}_\odot)$	$m_2\,(\mathrm{M}_\odot)$	$\mathcal{M}\,(\mathrm{M}_\odot)$	$q_{\min}$	ζ	Δc_α	N_{subbanks}	d_{subbanks}	$L_{\mathrm{max,subbanks}}$	$N_{\mathrm{templates}}$
`BNS 0`			< 1.1				1	2	777.0	48 806
`BNS 1`	$(1, 3)$	$(1, 3)$	$(1.1, 1.3)$	—	0.05	0.55	1	2	434.3	23 856
`BNS 2`			> 1.3				1	2	824.6	43 781
`NSBH 0`			< 3				1	4	753.4	84 641
`NSBH 1`	$(3, 100)$	$(1, 3)$	$(3, 6)$	1/50	0.05	0.5	2	6, 6	259.5, 166.8	85 149
`NSBH 2`			> 6				3	5, 4, 4	87.5, 61.2, 9.4	15 628
`BBH 0`			< 5			0.55	1	3	270.6	8246
`BBH 1`			$(5, 10)$			0.55	2	4, 4	113.7, 50.0	4277
`BBH 2`	$(3, 100)$	$(3, 100)$	$(10, 20)$	1/18	0.05	0.5	3	3, 4, 3	41.5, 33.5, 10.3	1607
`BBH 3`			$(20, 40)$			0.45	3	2, 2, 2	11.7, 10.8, 4.9	225
`BBH 4`			> 40			0.35	5	2, 2, 2, 1, 1	2.9, 2.0, 1.1, 0.7, 0.5	46
Total										316 262

Table 2.1: Summary of the parameters of the template banks. Columns 2 to 5 describe the bounds of physical parameter space that each bank is designed to cover. ζ is a tunable fudge parameter that controls the tolerance for removing nonphysical grid points. Δc_α is the grid spacing that we chose for each bank. N_{subbanks} is the resulting number of subbanks in each bank. d_{subbanks} and $L_{\mathrm{max,subbanks}}$ are the dimensionalities of each subbank (sorted by increasing mean total mass) and the size of their largest dimension, respectively. $N_{\mathrm{templates}}$ is the total number of templates in each bank.

banks. The banks differ greatly in size, which justifies the division of the search space into multiple banks.

For each subbank, we apply the procedure outlined in §2.2 to define a set of basis phase functions that generate a linear space and obtain the projections of the input waveforms onto that space. These are shown in Fig. 2.4 for the example case of the `BBH 0` bank, with the points color-coded by their chirp mass. The first three dimensions capture practically all the diversity of the input waveforms. Also note the large differences in size from the leading dimension to the sub-leading ones. The number of cycles, proportional to $\mathcal{M}^{-5/3}$, is the best-measured parameter and thus should approximately correspond to the coefficient of the leading dimension (Cutler and Flanagan, 1994; Dhurandhar and Sathyaprakash, 1994). Indeed, this is observed in Fig. 2.4, confirming that the decomposition is working as expected.

Next, we choose a grid spacing Δc_α common to all dimensions and define a rectangular grid in component space as follows. We force the point $\boldsymbol{c} = \boldsymbol{0}$ to be a grid point, because the SVD typically aligns the highest density regions (where the input physical waveforms tend to be) with the axes. Along each dimension, we add uniformly-spaced points until the whole range spanned by the input waveforms is covered. We allow the spacing to slightly decrease so that the most extreme input component is half the grid spacing away from the most extreme grid point. We do this for each dimension and in the positive and negative directions separately. Finally, not all the points of the rectangular grid describe physically viable waveforms. We only keep the templates that are

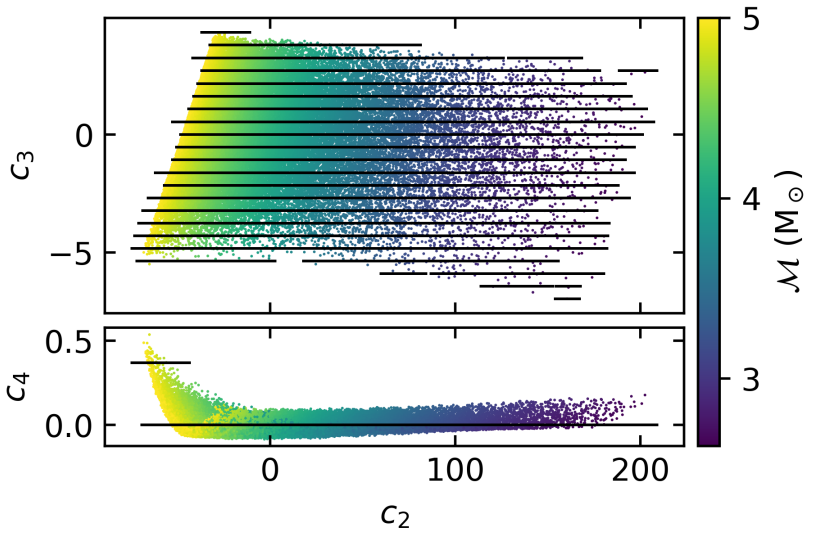


Figure 2.4: Our three-dimensional `BBH 0` template bank in the component space (black dots, which at this scale appear as lines), projected along two different axes. Underlaid are the input waveforms used to build the bank, projected according to Eq. (2.21) and color-coded by their chirp mass.

close to at least one input waveform, with the following criterion. For every input waveform set of components, we keep the closest grid point and a patch of the grid around it, with size equal to the corresponding dimension times a tunable fudge factor $\zeta \sim 0.1$.

Indeed, as Fig. 2.4 shows, the input physical waveforms do not fill the entire rectangular volume but are distributed within some irregularly shaped region. Furthermore, the density of input waveforms is low in the low-$\mathcal{M}$ region, where the waveforms have more wave cycles in band and hence are mutually more distinguishable. Holes can be produced in the physically viable region if the fudge factor ζ is too small, and there is an excess of unphysical templates if ζ is too large. We choose the Δc_α and ζ parameters such that we achieve a good balance between economic template bank size and high bank effectualness. The values chosen for each bank are reported in Table 2.1.

In Table 2.1 we observe a general trend with the mass: the banks for lighter mergers tend to have fewer subbanks and the first dimension spans a wider range. By comparison, the banks for heavier mergers have more subbanks, with smaller dimensions. The increase in the number of subbanks for heavier mergers is caused by the cutoff frequency falling in the band, which increases the variety of amplitude profiles.

There are interesting implications of the number of dimensions and their size for parameter estimation. Given an astrophysical signal, in the limit of high SNR ρ, the parameter likelihood is approximately given by $P(d \mid \boldsymbol{p}) \propto \exp\left(\rho^2|z|^2/2\right)$, where $z = (h(\boldsymbol{p}_*) \mid h(\boldsymbol{p}))$ is the complex match of $\boldsymbol{p}$ to the best-fit parameters $\boldsymbol{p}_*$ (Roulet and Zaldarriaga, 2019). By virtue of Eq. (2.15), this means

that the likelihood is approximately an isotropic Gaussian in terms of the c_α coordinates, with a width $\sim 1/\rho$. The number of dimensions can therefore be interpreted as the number of independent parameters that can be measured, and the size of each dimension as the relative precision that can be obtained for a fixed SNR (with the caveat that we have restricted the frequency range; for example, information about the tidal deformability comes from frequencies higher than our 512 Hz cutoff).

For example, for BNS (and effectively for light BBH) the banks have two dimensions, with a large first dimension well correlated with the chirp mass (Fig. 2.4). The two measurable parameters are the chirp mass, which indeed can be measured to much higher precision than for heavy systems, and a combination of the mass ratio and effective spin which can be measured with a lower precision. These are the leading contributions to the phase evolution as can be understood from the post-Newtonian expansion.

An important advantage of our geometric coordinates is that they are well suited for a two-step search that effectively achieves a smaller grid spacing at reduced computational cost. We realize this by refining the template grid on demand around all triggers that exceed an appropriately lowered SNR threshold (Venumadhav et al., 2019; Gadre et al., 2019). During the search, we first use a coarse grid, and refine every trigger using neighboring templates from a denser grid that has half the spacing along each dimension. The fact that the distance between c_α components translates directly to mismatch (Eq. (2.15)) makes this method straightforward to implement.

To characterize the effectualness of the bank at recovering the target physical signals, we generate a set of 10^4 random "test waveforms" within the parameter range of each bank, using the same approximant with which the input waveforms were generated. We choose the parameters from a distribution that is uniform in the component masses m_1, m_2 and aligned spins χ_1, χ_2. In principle, we would have to match each test waveform against every waveform in the bank to look for the best match. To save computational effort, we select a candidate best-match based on the approximate metric Eq. (2.15) by extracting the phase of the test waveform $\psi^{(i)}(f)$, projecting it onto the linear space, $c_\alpha^{(i)} = \langle \psi^{(i)} - \overline{\psi}, \psi_\alpha \rangle$, and finding the closest grid point with respect to the Euclidean metric (2.15). Since *a priori* we do not know which subbank best describes the test waveform, we pick the best candidate from each subbank and compute the match with all. The best match with our reduced set of candidates is a lower bound on the best match over all the waveforms in the bank. Rather than using Eq. (2.4) directly, we compute the match by following the detection strategy described in Venumadhav et al. (2019): we account for the finite time resolution of the Fourier transform by downsampling the waveforms to 512 Hz and sinc-interpolating the matched-filter output twice. We show the result of this test in Fig. 2.5, in terms of the cumulative fraction of

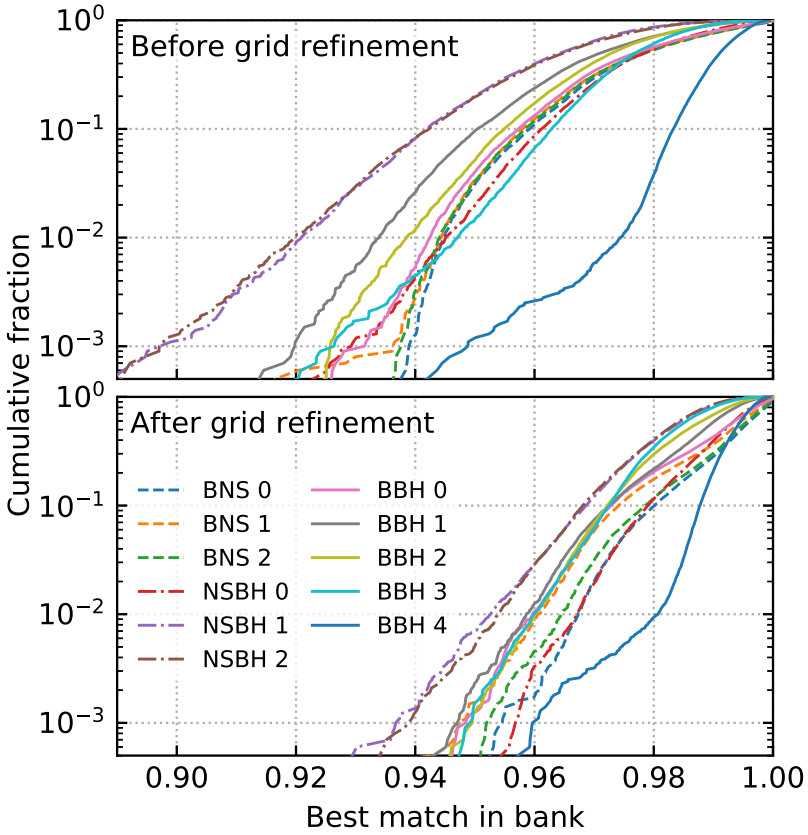


Figure 2.5: Effectualness of our template banks, tested on random waveforms drawn from a distribution uniform in individual masses and aligned spins. The vertical axis shows the fraction of the random trials that do not achieve a given match in the bank.

the matches with each bank before and after applying the grid refinement, which we use to assess the collection threshold on the coarse grid and the effectualness achieved for each bank, respectively. We find that depending on the bank 99% of the templates have a match higher than 0.95 to 0.98.

2.4 Conclusions

We have developed a general and computationally efficient geometric placement algorithm to construct high-effectualness template banks for detecting gravitational waves from compact binary mergers. We have constructed a basis of functions that generate a linear space of phase profiles on which the mismatch metric is Euclidean. For the purpose of signal detection, we shift the focus away from physical parameters to the linear coefficients for the basis phase profiles. We identified which components carry the largest amount of information about physical waveforms and what is the minimal set required to guarantee a desired match. The basis functions can be determined from a set of input waveforms whose size is small compared to that of the bank. The basis functions

can be generated with any frequency-domain waveform model. The resolution of the bank can be decided independently after the basis functions have been found; in particular, it can be increased arbitrarily at negligible computational cost since no further evaluations of the physical waveform approximants need to be done. Our algorithm guarantees that within each of the few subbanks that make up one template bank, all templates share the same amplitude profile, a property that is critical for the correction of the power-spectral-density drift in signal processing.

We have applied our algorithm to the construction of a collection of eleven template banks that together cover the parameter space associated to stellar-mass compact binary mergers with aligned spins. We find the effectualness and total number of templates to be comparable to the ones obtained by other algorithms in the literature (Canton and Harry, 2017; Roy et al., 2019); detailed comparisons are difficult due to the different parameter spaces targeted in various works. We note that our template bank includes rapidly spinning neutron stars, which to date have not been searched for in the gravitational wave data. We implement a two-step search with a coarse grid that we refine around triggers at the time of search, a task for which our new formalism is ideally suited. This is an important step to reduce the number of templates while preserving a high effectualness.

Looking forward, an accurate and fast interpolation from physical parameters to the c_α component space would be extremely useful for rapid parameter estimation. First, because waveforms can be generated at negligible computational cost once the components are known. At least in cases where analytical waveform models are not valid, waveform generation dominates the computational cost of parameter estimation. Moreover, the likelihood would look close to an isotropic Gaussian in terms of the c_α coordinates due to orthonormality, making them a suitable choice from the data analysis perspective. Other natural extensions of the work presented here are to include the effects of precession, due to misalignment between the spins and the orbital angular momentum, and eccentricity. These are deferred for future work. The inclusion of eccentricity is currently limited by the availability of robust public waveform generation codes.

The template bank described here is available at `https://github.com/jroulet/template_bank`.

Chapter 3

Early Constraints on Binary Black Hole Populations

This Chapter has been adapted from Roulet and Zaldarriaga (2019).

We reanalyse the LIGO–Virgo strain data of the 10 binary black hole mergers reported to date and compute the likelihood function in terms of chirp mass, mass ratio and effective spin. We discuss the strong degeneracy between mass ratio and spin for the three lighter events. We use this likelihood and an estimate of the horizon volume as a function of intrinsic parameters to constrain the properties of the population of merging binary black holes. The data disfavour large spins. Typical spins are constrained to $\overline{a} \lesssim 0.4$, even if the underlying population has randomly-oriented spins. For aligned spins the constraints are tighter, with typical spins required to be around $\overline{a} \sim 0.1$ and have comparable dispersion. We detect no statistically significant tendency towards a positive average spin in the direction of the orbital angular momentum. We put an upper limit on the fraction of systems where the secondary could have been tidally locked prior to the formation of the black holes (corresponding to merger times shorter than 10^8 years) $f \lesssim 0.3$. Four events are consistent with having a maximally-spinning secondary, although one only marginally. We confirm previous findings that there is a hint of a cutoff at high mass. The data favour distributions of mass ratios with an average $\overline{q} \gtrsim 0.7$.

3.1 Introduction

Binary black holes (BBH) have been observed for the first time with the recent advent of gravitational wave (GW) observatories (Abbott et al., 2016a,b,c, 2017b,c,d, 2019a). The astrophysical origin of these systems remains a major open question. Potential formation channels that have been proposed in the literature include isolated binary evolution through a common envelope phase, a chemically homogeneous evolution in a tidally locked binary, the dynamical formation in dense stellar environments such as globular clusters or in triple systems and the formation in galactic nuclear disks assisted by the presence of gas (references to various scenarios and how they fare in comparison with the LIGO data can be found in Abbott et al. (2016c)). Although the gravitational wave data are yet insufficient to decide between these scenarios, constraints on specific models can already start to be set (see for example results in Abbott et al. (2016c); Vitale et al. (2017b); Talbot and Thrane (2017); Farr et al. (2017); Hotekezaka and Piran (2017); Farr et al. (2018); Abbott et al. (2019b)).

To use the LIGO–Virgo events to constrain the properties of the population of merging black holes we need the likelihood of individual events as a function of the parameters of the binary. Although constraints on individual parameters have been reported by the LIGO and Virgo collaborations (LVC), the full likelihoods have not been provided to the community yet (posterior samples have been recently released). These are necessary in order to properly include the correlations between parameters, which as we will see are substantial. Thus in this paper we set out to reanalyse the LIGO–Virgo strain data (under some simplifying assumptions) to obtain likelihoods for the events and then use those to set constraints on the parameters of population models for the BBHs. The attempts to put constraints on the populations already presented in the literature were limited by the lack of the individual event likelihoods, so approximate constraints were based on properties of the one-dimensional posteriors gathered from the LVC figures and papers.

Neglecting center-of-mass acceleration (e.g. due to a third body), BBH merger events are completely characterized by 17 parameters (two masses, six spin components, two angles for the orbit orientation, two for the sky location, distance, merger phase and time, orbital eccentricity and pericenter angle). However, the following key observations allow us to significantly reduce the number of parameters considered.

First, not all the parameters are informative of the population properties. Population models generally predict a homogeneous and isotropic GW-source distribution on scales resolvable by GW detectors, so the source location is irrelevant to distinguish among these models. Similarly, the source orientation and the time and phase of merger are uniformly distributed. Moreover, the signal

dependence on these parameters is well understood, so the likelihood can be marginalized over them.

Second, not all the parameters can be constrained by the data at the current sensitivity levels. Only one of the spin combinations, $\chi_{\rm eff}$, is relatively well measured (see for example Vitale et al. (2014, 2017a)). It is defined by

$$\chi_{\rm eff} = \frac{m_1\vec{a}_1 + m_2\vec{a}_2}{m_1 + m_2} \cdot \hat{L}, \tag{3.1}$$

where $\vec{a}_i = c\vec{S}_i/Gm_i^2$ and $\hat{L}$ is the direction of the orbital angular momentum. Spin components orthogonal to $\chi_{\rm eff}$ are largely unconstrained. Thus, including these parameters in the analysis increases the computational cost without significantly changing the results. Moreover, the waveform templates currently used for detection and parameter estimation have no eccentricity, so LIGO is not able to measure it. We will assume no eccentricity as well.

Third, astrophysical models for the populations are quite crude and thus small shifts in the parameters or error bars are not likely to change the astrophysical conclusions one might draw. At this stage one is interested in more qualitative questions such as whether the BHs are spinning fast, whether the spins tend to be aligned normal to the orbital plane or what the range of masses of the BBHs is.

As a result of these considerations, there is a clear hierarchy in the parameters based on how much they can constrain BBH population models. $\chi_{\rm eff}$ and the two masses are measurable and have distributions dependent on the models (Mandel and O'Shaughnessy, 2010; Rodriguez et al., 2016b; Zevin et al., 2017; Stevenson et al., 2017a). The other spin components and the eccentricity, while dependent on the population model, are still poorly constrained by the data, so we henceforth ignore them. The various angles, phase, time of merger and distance are uninformative of the population, so they are nuisance parameters for our purposes. They have a known effect on the signal, so we marginalize over them in §3.2.

We will parametrize the two masses in terms of the chirp mass $\mathcal{M}$ and the mass ratio q, given by

$$\begin{aligned} \mathcal{M} &= \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}} \\ q &= \frac{m_2}{m_1} \leq 1, \end{aligned} \tag{3.2}$$

because $\mathcal{M}, q$ are generally less correlated than m_1, m_2, as these are the combinations that respectively enter the GW waveform at the leading and the next-to-leading order in the post-Newtonian expansion. To implement our restriction to a single spin variable, we will use the prescription that

the spins are aligned with the orbital angular momentum, and that $\chi_1 = \chi_2 = \chi_{\rm eff}$, where $\chi_i = \vec{a}_i \cdot \hat{L}$. At the current level of sensitivity the exact way one distributes the spin to obtain a given $\chi_{\rm eff}$ does not affect the parameter constraints in a meaningful way.

We will denote the informative parameters by $\boldsymbol{p} = (\mathcal{M}, q, \chi_{\rm eff})$. Unless otherwise stated, we will refer to the detector-frame mass, whose value is redshifted from the source-frame mass by $\mathcal{M} = (1+z)\mathcal{M}_{\rm source}$. Both q and $\chi_{\rm eff}$ are independent of redshift.

3.2 Single-event likelihood

In this section we describe our computation of the BBH-parameter likelihood of a GW event and its analytical marginalization over the nuisance parameters. What we do is rather standard but we detail our procedure so that we can report all our simplifications. We first analyse the case of a single detector.

We will define the noise-weighted inner product between two functions in the frequency domain

$$\langle x \mid y \rangle = 4\,\mathrm{Re} \int_0^\infty \frac{\tilde{x}^*(f)\tilde{y}(f)}{\sigma^2(f)} \mathrm{d}f, \tag{3.3}$$

where $\sigma^2(f)$ is the one-sided power spectral density (PSD) of the detector noise and the tildes indicate Fourier transforms. Under the assumption that the noise is additive, stationary and Gaussian, the single-detector likelihood $\mathcal{L} = P(d \mid h)$ that the data d have been produced by a model GW signal h is

$$\log \mathcal{L} = -\frac{1}{2}\langle d - h \mid d - h \rangle. \tag{3.4}$$

We have access to both d and h, because the strain data for all reported events and the approximants for generating template waveforms have been released by the Gravitational Wave Open Science Center (GWOSC) (Vallisneri et al., 2015). We use two different approximants, `SEOBNRv4_ROM`, based on the effective-one-body formalism (Bohé et al., 2017), and `IMRPhenomD`, based on a phenomenological approach (Khan et al., 2016), as a robustness test of our results. We estimate the PSD using the `PyCBC` (Biwer et al., 2018) implementation of the median-mean spectrum described in Allen et al. (2012) on a segment of 32 s of data centered around each event. As recommended by the GWOSC, we use a lower frequency cutoff of 10 Hz except for GW170608, which we cut at 20 Hz for Livingston and 30 Hz for Hanford.

We decompose the modeled signal into the form

$$\tilde{h}(f;\boldsymbol{p},a_0,t_0,\phi_0) = a_0 \mathrm{e}^{\mathrm{i}(2\pi f t_0 - \phi_0)} \frac{\tilde{h}_0(f;\boldsymbol{p})}{\sqrt{\langle h_0 \mid h_0\rangle}}. \tag{3.5}$$

Here $\tilde{h}_0(f;\boldsymbol{p})$ is the waveform template, that depends on the set of physical parameters $\boldsymbol{p} = (\mathcal{M}, q, \chi_{\mathrm{eff}})$ that are intrinsic to the binary. We work in the quadrupole radiation approximation, i.e. that the dominant spherical harmonics of the emitted GW are $(l,m) = (2,\pm 2)$. The shape of h_0 is then independent of the nuisance parameters, which only enter through a_0, ϕ_0, t_0. The inclusion of other multipole moments would introduce a dependence of the waveform shape on the orientation of the binary (Cotesta et al., 2018). We divide h_0 by the norm to eliminate its arbitrary normalization. We may then compute h_0 at any fiducial configuration, e.g. directly above the detector at a 1 Mpc distance, with a face-on alignment. t_0 and ϕ_0 are the time and phase of the merger as seen in the detector. a_0 is the amplitude, which can be interpreted as the expectation value for the signal-to-noise ratio (SNR) with which a signal identical to h would be measured in the detector, given by $\sqrt{\langle h \mid h\rangle}$. For multiple detectors, each one will have different values of a_0, t_0, ϕ_0 that are related by the various angles between source and detectors.

Combining (3.4) and (3.5) we obtain

$$\begin{aligned} \log \mathcal{L} &= -\frac{1}{2}\left(\langle d \mid d\rangle - 2\langle d \mid h\rangle + \langle h \mid h\rangle\right) \\ &= -\frac{1}{2}\left(\langle d \mid d\rangle - 2a_0|z|\cos(\phi_0 - \arg z) + a_0^2\right), \end{aligned} \tag{3.6}$$

where $z(\boldsymbol{p}, t_0)$ is the complex matched filter output (Allen et al., 2012):

$$z(\boldsymbol{p},t_0) = \frac{4}{\sqrt{\langle h_0 \mid h_0\rangle}} \int_0^\infty \frac{\tilde{d}^*(f)\tilde{h}_0(f;\boldsymbol{p})}{\sigma^2(f)} \mathrm{e}^{\mathrm{i}2\pi f t_0} \mathrm{d}f. \tag{3.7}$$

Note that, for any set of parameters $\boldsymbol{p}$, we can compute z for all t_0 with a single Fast Fourier Transform.

3.2.1 Likelihood marginalization

We will now marginalize the likelihood (3.6) over the nuisance parameters a_0, ϕ_0, t_0, since their priors do not depend on the BBH population model, and we will keep the dependence on $\boldsymbol{p}$. Our approach will be to do the marginalization subject to the condition $\mathcal{D}$ that the event has been detected, that

is, we will define

$$\begin{aligned}\overline{\mathcal{L}}(\boldsymbol{p}) &\equiv P(d \mid \boldsymbol{p}, \mathcal{D}) \\ &= \int \mathrm{d}a_0 \mathrm{d}\phi_0 \mathrm{d}t_0 P_{\mathrm{prior}}(a_0, \phi_0, t_0 \mid \boldsymbol{p}, \mathcal{D}) \mathcal{L}(a_0, \phi_0, t_0, \boldsymbol{p} \mid \mathcal{D}) \\ &= \int \mathrm{d}a_0 \mathrm{d}\phi_0 \mathrm{d}t_0 P_{\mathrm{prior}}(a_0, \phi_0, t_0 \mid \mathcal{D}) \mathcal{L}(a_0, \phi_0, t_0, \boldsymbol{p}). \end{aligned} \tag{3.8}$$

The last equality follows because, as we will show, the prior for a_0, ϕ_0, t_0 does not depend on $\boldsymbol{p}$ once conditioned to detectability; and the likelihood $\mathcal{L}$ is suppressed for combinations of parameters that yield undetectable signals, since the data under consideration correspond to detections and detectability is a property of the data only. $P(d \mid \boldsymbol{p}, \mathcal{D})$ differs from $P(d \mid \boldsymbol{p})$ in that it excludes the selection bias of the detector, whose sensitivity depends on $\boldsymbol{p}$. Indeed, for data corresponding to detections

$$P(d \mid \boldsymbol{p}) = P(d \mid \mathcal{D}, \boldsymbol{p}) P(\mathcal{D} \mid \boldsymbol{p}), \tag{3.9}$$

where $P(\mathcal{D} \mid \boldsymbol{p})$ is the selection bias (see Appendix 3.A). By imposing the detection condition $\mathcal{D}$ we are taking the observer's point of view, where the events analysed are conditioned to detection and the selection bias enters in the form of a prior for $\boldsymbol{p}$. We show the equivalence to the alternative, physical approach usually taken in the literature (e.g. Fishbach and Holz (2017); Mandel et al. (2019)) in Appendix 3.A.

We compute the prior in Eq. (3.8) as follows. ϕ_0 and t_0 have uniform priors. a_0 is proportional to D^{-1}, where D is the luminosity distance to the event. At low redshifts, the prior for the distance is $P_{\mathrm{prior}}(D) \propto D^2$, so demanding $P_{\mathrm{prior}}(D)\mathrm{d}D = P_{\mathrm{prior}}(a_0)\mathrm{d}a_0$ yields $P_{\mathrm{prior}}(a_0) \propto a_0^{-4}$. For the event to have been a detection, a_0 must exceed a certain threshold value, for which we take a conservative (low) value of $a_{0,\mathrm{min}} = 9$ (see fig. 9 of Abbott et al. (2016c)).

The priors for the nuisance parameters are then

$$P_{\mathrm{prior}}(\phi_0) = \frac{1}{2\pi} \tag{3.10}$$

$$P_{\mathrm{prior}}(t_0) = \frac{1}{T} \tag{3.11}$$

$$P_{\mathrm{prior}}(a_0) = \frac{3a_{0,\mathrm{min}}^3}{a_0^4}, \qquad a_0 > a_{0,\mathrm{min}} \tag{3.12}$$

where T is the duration of the data being analysed and the priors have been normalized to integrate

to 1 over their domains, as required by the detection condition.

Using (3.6), (3.10) and (3.12), the marginalization of the likelihood over ϕ_0 and a_0 yields:

$$\int_{a_{0,\min}}^{\infty} \mathrm{d}a_0 P_{\mathrm{prior}}(a_0) \int_0^{2\pi} \frac{\mathrm{d}\phi_0}{2\pi} \mathcal{L} = 3a_{0,\min}^3 \mathrm{e}^{-\frac{1}{2}\langle d|d\rangle} I(|z|), \tag{3.13}$$

where we have defined

$$I(|z|) = \int_{a_{0,\min}}^{\infty} \frac{I_0(a_0|z|)\mathrm{e}^{-\frac{1}{2}a_0^2}}{a_0^4} \mathrm{d}a_0. \tag{3.14}$$

Here, I_0 is the modified Bessel function of the first kind of order zero. In practice, we ignore the constant factor in front of $I(|z|)$ in (3.13) as it does not depend on the parameters. To implement the computation of $I(|z|)$ efficiently, we tabulate its more smoothly varying logarithm for several values of $|z|$ and interpolate in between.

Finally, we can further marginalize Eq. (3.13) with respect to t_0 using (3.7) and (3.11), computing

$$\int_{a_{0,\min}}^{\infty} \mathrm{d}a_0 P_{\mathrm{prior}}(a_0) \int_0^{2\pi} \frac{\mathrm{d}\phi_0}{2\pi} \int_0^T \frac{\mathrm{d}t_0}{T} \mathcal{L}(\boldsymbol{p}, a_0, \phi_0, t_0) \propto \int_0^T I(|z(\boldsymbol{p}, t_0)|)\mathrm{d}t_0 \tag{3.15}$$

by numerical quadrature.

Eq. (3.15) gives the one-detector likelihood marginalized over the nuisance parameters a_0, ϕ_0, t_0, assuming non-precessing spins and that the dominant mode of GW emission is $(l, m) = (2, \pm 2)$. For the case of multiple detectors, the total likelihood is the product of all the one-detector likelihoods, but the marginalization should be made over the source parameters (location, orientation, phase and time of coalescence), that correlate the values of a_0, ϕ_0, t_0 observed at each detector.

The case of LIGO is particularly extreme because by design, the two detectors at Hanford and Livingston have the same orientation to a good approximation (plus a 90° rotation in the plane of the detector). Under the approximation that the two detectors are co-aligned, a signal must have the same phase ϕ_0 and strain amplitude $A_0 \equiv a_0\sqrt{\langle h_0 \mid h_0\rangle}$ in both detectors. The arrival times, however, can be different depending on the location of the source, so these must be marginalized over separately. The time delay between detectors is $\delta t = t_d \cos\theta$, where $t_d = 10.012\,\mathrm{ms}$ is the GW travel time between sites and θ is the angle between the source and a line passing through both detectors. Although an isotropic distribution of sources is uniform in $\cos\theta$ and thus in δt, the antenna pattern of the detectors induces a selection bias because sources above or below the plane of the detectors are more likely to be detected (Sathyaprakash and Schutz, 2009; Chen et al., 2017). The resulting

prior for δt is well fit by (Cornish et al., 2017)

$$P_{\rm prior}(\delta t) \propto 1 - \left(\frac{\delta t}{10.65\,{\rm ms}}\right)^2, \qquad |\delta t| \leq t_d. \tag{3.16}$$

Since both detectors measure the same polarization, for any given δt we can combine the Hanford and Livingston streams of data coherently into a single "combined channel", $d_{\rm LIGO}$, in a way that minimises the relative variance:

$$\tilde{d}_{\rm LIGO}(f;\delta t) = \sigma^2_{\rm LIGO}(f)\left(\frac{\tilde{d}_{\rm H}(f)}{\sigma^2_{\rm H}(f)} - {\rm e}^{-{\rm i}2\pi f\delta t}\frac{\tilde{d}_{\rm L}(f)}{\sigma^2_{\rm L}(f)}\right), \tag{3.17}$$

where H and L refer to Hanford and Livingston, and $\sigma^2_{\rm LIGO} = (\sigma^{-2}_{\rm H} + \sigma^{-2}_{\rm L})^{-1}$ is the noise PSD of the combined channel. The minus sign accounts for the relative 90° rotation between the two LIGO sites. Since the true value of δt is not known, this parameter has to be marginalized over. With the combined channel, we can use the single-detector formulas to get the marginalized likelihood of the LIGO network. From (3.17) and (3.7) we obtain

$$z_{\rm LIGO}(\boldsymbol{p}, t_0, \delta t) = f^{1/2}_{\rm H} z_{\rm H}(\boldsymbol{p}, t_0) - f^{1/2}_{\rm L} z_{\rm L}(\boldsymbol{p}, t_0 + \delta t), \tag{3.18}$$

where

$$f_\alpha = \frac{\langle h_0 \mid h_0\rangle_\alpha}{\langle h_0 \mid h_0\rangle_{\rm H} + \langle h_0 \mid h_0\rangle_{\rm L}}, \qquad \alpha \in \{{\rm H}, {\rm L}\}. \tag{3.19}$$

We finally obtain the LIGO-network marginalized likelihood using (3.15), (3.16) and (3.18):

$$\overline{\mathcal{L}}_{\rm LIGO}(\boldsymbol{p}) \propto \int_0^T {\rm d}t_0 \int_{-t_d}^{t_d} {\rm d}\delta t P_{\rm prior}(\delta t) I(|z_{\rm LIGO}(\boldsymbol{p}, t_0, \delta t)|), \tag{3.20}$$

which we evaluate by quadrature. Note that $a_{0,\rm min}$ is now interpreted as the detection threshold on the LIGO-network SNR.

Although in this analysis we have been careful to analyse the data of both detectors coherently we have checked that this is largely unimportant for the constraints on the populations we obtain. One gets effectively the same constraints if one treats each detector independently using Eq. (3.15) and combines them incoherently, $\log \overline{\mathcal{L}}_{\rm LIGO}(\boldsymbol{p}) \approx \log \overline{\mathcal{L}}_{\rm H}(\boldsymbol{p}) + \log \overline{\mathcal{L}}_{\rm L}(\boldsymbol{p})$. The change in the likelihood distribution compared to its width was on the few-percent level or smaller in all cases.

The first five detections reported to date, as well as the last one, are LIGO-only, for which (3.20) is accurate. The remaining four events were also observed in Virgo. Since Virgo is not co-aligned with

LIGO, we cannot apply the treatment above. Instead, we make the approximation that a_0, ϕ_0, t_0 are uncorrelated between LIGO and Virgo. This amounts to ignore the fact that we know the relative orientations, locations and timing between those detectors. Since we are discarding information, this approximation will increase the uncertainties in $\boldsymbol{p}$ without biasing the maximum-likelihood values. For the three-detector events, then, we use

$$\overline{\mathcal{L}}_{\mathrm{HLV}}(\boldsymbol{p}) \propto \overline{\mathcal{L}}_{\mathrm{LIGO}}(\boldsymbol{p}) \int_0^T I\left(|z_{\mathrm{V}}(\boldsymbol{p}, t_0)|\right) \mathrm{d}t_0, \tag{3.21}$$

where V stands for Virgo. As we explained before, treating the detectors independently is sufficient for our purposes even for LIGO, so this is an excellent approximation in the context of our simplified analysis.

Since we have kept only three parameters, it is practical to evaluate the likelihood (3.21) over a grid in $\boldsymbol{p} = (\mathcal{M}, q, \chi_{\mathrm{eff}})$ for each event. We use a regular grid of 64^3 points, centered around the values reported by the LVC and with an extent of twice the reported uncertainties (subject to the bounds $0 < q \leq 1, -1 \leq \chi_{\mathrm{eff}} \leq 1$). In all cases we verify on random parameter values that interpolating the likelihood from the grid has good agreement with the actual computation. By using a grid, we have to compute the likelihood only once, and we are able to apply any prior easily a posteriori. This is a key requirement for model inference, since the prior for the parameter values depends on the population model. Working in a low-dimensional parameter space enables us to circumvent the need for Monte Carlo Markov chains.

3.2.2 LIGO–Virgo reported binary black holes

The single-event marginalized likelihood computed in this way is shown as a function of the parameters in Fig. 3.1. We obtained very similar results using the `SEOBNRv4_ROM` and `IMRPhenomD` approximants, so we only show the results for `SEOBNRv4_ROM`. The likelihood can be interpreted as the posterior distribution that would arise form a uniform prior in $\mathcal{M}, q, \chi_{\mathrm{eff}}$. In Appendix 3.B we show each event in greater detail, and apply the LIGO prior as a check of our pipeline.

In the middle panel of Fig. 3.1 we superimposed the masses and spins of the black holes in X-ray binaries (XRBs) as reported in McClintock et al. (2013). The variables for XRBs (m, a) and BBHs $(\mathcal{M}, \chi_{\mathrm{eff}})$ are different, so care has to be taken when comparing them. By definition (Eq. (3.2)), the individual masses of the BBHs are guaranteed to satisfy $m_2 \leq 2^{1/5}\mathcal{M} \leq m_1$, with $2^{1/5} \approx 1.15$. We recall that χ_{eff} is a mass-weighted average of the spin components of the two black holes parallel to the orbital angular momentum. We comment on the XRBs in §3.4.1.

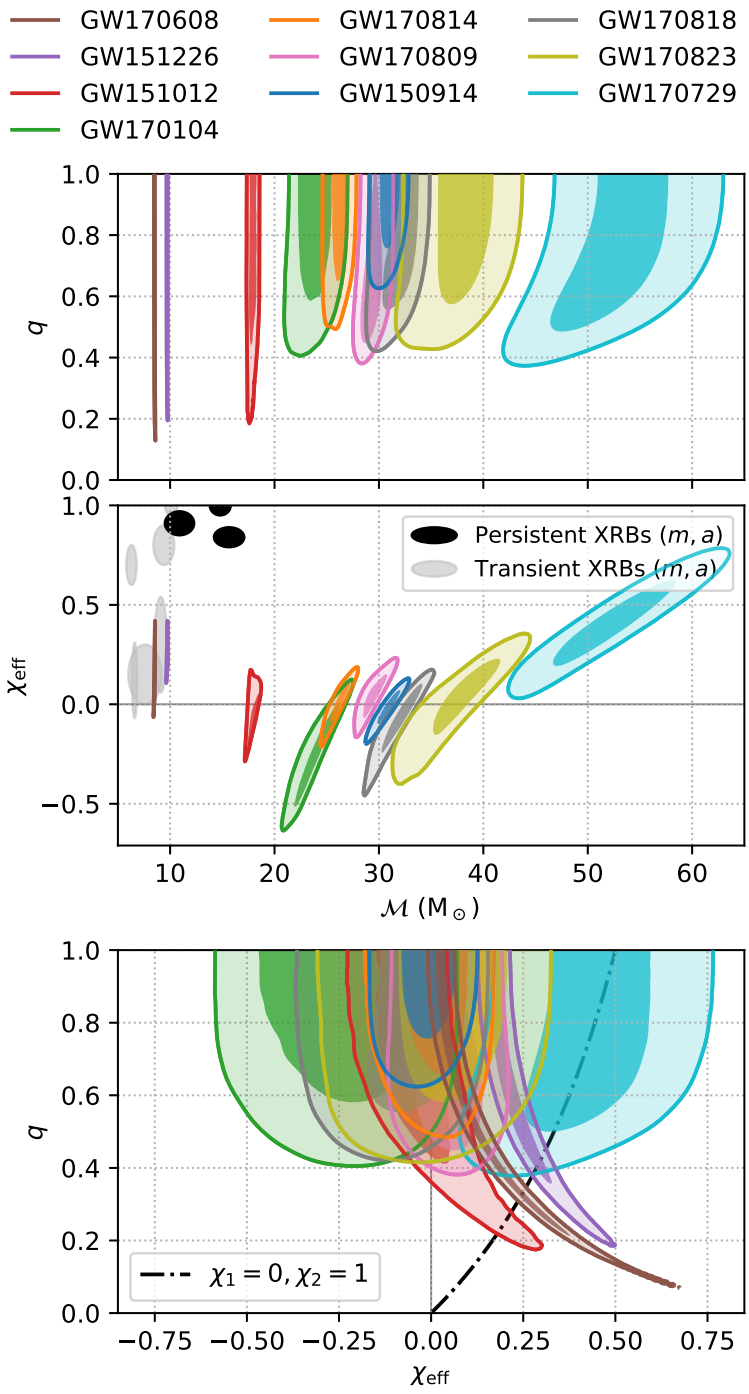


Figure 3.1: Marginalized likelihood contours enclosing 50% and 90% of the distribution for each of the six events reported to date. In the middle panel, the persistent and transient X-ray binaries reported in McClintock et al. (2013) are shown by ellipses, whose position in the plane represents the black hole mass m and its dimensionless spin a. In the bottom panel, the case where the aligned spins of the black holes are $\chi_1 = 0, \chi_2 = 1$ is shown by a dashed-dotted line, as a proxy for what the outcome of a tidally-locked-secondary progenitor would be (see §3.4.1).

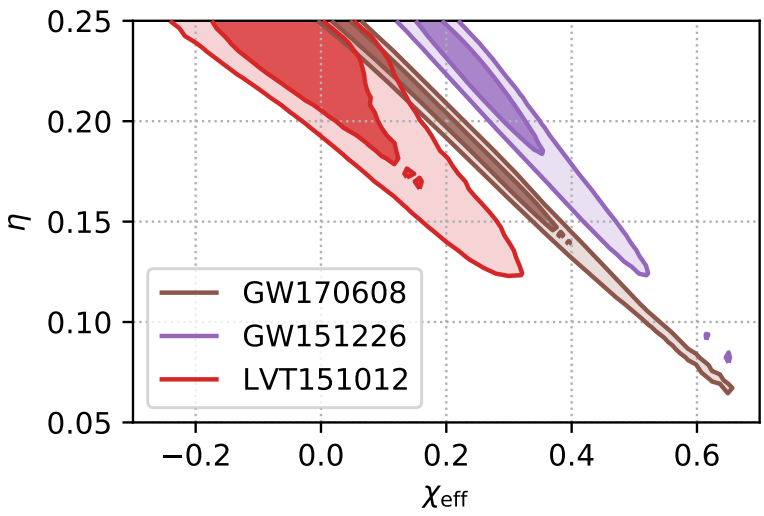


Figure 3.2: Parameter likelihood for the three lighter likely events, marginalized over $\mathcal{M}$, as a function of $\eta = q/(1+q)^2$ and $\chi_{\rm eff}$. At low mass, η and $\chi_{\rm eff}$ are degenerate.

It is interesting to note that there are clear degeneracies between the parameters and that those degeneracies change with the mass of the system. At low mass there is a strong degeneracy between mass ratio and effective spin. This was of course expected (for an early discussion see Cutler and Flanagan (1994)). Low mass binaries merge at a higher frequency and thus the detector is more sensitive to the inspiral, where the post-Newtonian (PN) expansion is accurate. The leading PN corrections including spin are approximately degenerate with the leading corrections including mass ratio. This correlation is simpler when expressed using $\chi_{\rm eff}$ and the symmetric mass-ratio $\eta = q/(1+q)^2$ as variables (Baird et al., 2013; Ng et al., 2018), as we show in Fig. 3.2 (compare to the bottom panel of Fig. 3.1). It is apparent that a linear combination of the two parameters is better constrained than each of them. We report this combination in Table 3.1.

At high mass we observe a different degeneracy, between the chirp mass and effective spin (middle panel of Fig. 3.1). In General Relativity, the mass can be scaled out of the problem as a time-scale. That is, a system with a smaller chirp mass and all other dimensionless parameters constant yields the same waveform, but sped up, or shifted to higher frequencies. A positive aligned spin can mimic this effect: it acts as an effective repulsive force between the BHs, which delays the plunge and makes them merge at a higher frequency (Campanelli et al., 2006). Since at high masses the chirp mass is not too well constrained, the degeneracy appears. Like before, a linear combination of $\mathcal{M}, \chi_{\rm eff}$ is better constrained than either parameter, which we also report on Table 3.1.

Thus it is more convenient to change parameter basis and use a different combination for the

Event	A	$\chi_{\rm eff} + A(\eta - 1/4)$
GW151226	2.8	$0.16^{+0.04}_{-0.03}$
GW151012	3	$-0.09^{+0.12}_{-0.1}$
GW170608	3.4	$0.02^{+0.04}_{-0.02}$
	B ($\mathrm{M}_\odot$)	$\mathcal{M} - B\chi_{\rm eff}$ ($\mathrm{M}_\odot$)
GW170104	8	$25.9^{+0.8}_{-1.0}$
GW170814	8	$26.1^{+0.6}_{-0.5}$
GW170809	9	$28.9^{+0.8}_{-0.6}$
GW150914	12	$31.0^{+0.6}_{-0.5}$
GW170818	11	$32.7^{+1.0}_{-0.7}$
GW170823	15	37^{+2}_{-2}
GW170729	27	41^{+3}_{-3}

Table 3.1: Best constrained linear combination of $\eta, \chi_{\rm eff}$ for the three lighter events, and of $\mathcal{M}, \chi_{\rm eff}$ for the seven heavier ones, with 90% confidence uncertainties. A and B are coefficients chosen to minimise the correlation.

heavy and light events. In the new basis the constraints can be better approximated by a simple Gaussian. Such approximation can provide a quick way to make estimates of the population parameters for astrophysical use.

3.3 Model inference

3.3.1 Universe- and detector-rates

We turn now to the task of constraining population models combining the data of all events. We will do that by introducing a new set of population-parameters $\boldsymbol{\lambda}$, that will depend on the specific model at hand and which we want to constrain. We make two remarks in that respect.

First, we note that what we can constrain is the rate at which BBHs with certain parameter values (masses, spins) merge, which is not necessarily proportional to the abundance of those systems, since their dynamics depend on the parameter values. Light binaries, for example, need to start closer in order to merge in less than the age of the Universe, since GW emission is less efficient than for heavy BBHs. Then, our definition of a model, parametrized by $\boldsymbol{\lambda}$, is given by the volumetric merger rate $R(\boldsymbol{p_s}, z_{\rm rs} \mid \boldsymbol{\lambda})$, that can depend on the cosmological redshift $z_{\rm rs}$ and the source-frame parameters $\boldsymbol{p_s} = (\mathcal{M}_{\rm s}, q, \chi_{\rm eff})$. We recall that $\mathcal{M}_{\rm s} = \mathcal{M}/(1 + z_{\rm rs})$, while $q, \chi_{\rm eff}$ are independent of redshift. R can explicitly depend on redshift if the merger rate depends on time. For example, star formation peaked at $z_{\rm rs} \sim 2$, so if the delays between black hole formation and merger are short compared to

the age of the Universe, the merger rate can be expected to be higher at large redshifts. Instead, we will assume the rate to be uniform in comoving volume, so that $R(\boldsymbol{p_s} \mid \boldsymbol{\lambda})\mathrm{d}\boldsymbol{p_s}$ is the rate at which BBHs with parameter values within $\mathrm{d}\boldsymbol{p_s}$ of $\boldsymbol{p_s}$ merge per unit comoving volume and source-frame time.

Second, the rate of events observed at the detector network is not proportional to the merger rate in the Universe, because of two effects. First, the strength of the GW signal, and thus the horizon distance up to which an event can be observed, again depends on the parameters, inducing a selection bias. And second, the detector- and source-frame masses differ because of cosmological redshift. There are additional redshift considerations because GW detectors are sensitive to a luminosity-volume, while we defined the rate per comoving-volume, and because the merger rate is redshifted. To incorporate these effects, we introduce the detection rate Λ, such that $\Lambda(\boldsymbol{p} \mid \boldsymbol{\lambda})\mathrm{d}\boldsymbol{p}$ is the rate of detection of events with detector-frame parameters $\boldsymbol{p}$ per unit time. It is related to the physical rate R by

$$\begin{aligned}\Lambda(\boldsymbol{p} \mid \boldsymbol{\lambda}) &= 4\pi \int \mathrm{d}\mathcal{M}_\mathrm{s} \int \mathrm{d}D_c D_c^2 \frac{R(\boldsymbol{p_s} \mid \boldsymbol{\lambda})}{1+z_\mathrm{rs}} \delta(\mathcal{M} - \mathcal{M}_\mathrm{s}(1+z_\mathrm{rs})) f(\boldsymbol{p}, D_L) \\ &= 4\pi \int \mathrm{d}D_L D_L^2 \frac{1}{(1+z_\mathrm{rs})^5} \left(1 - \frac{D_L}{1+z}\frac{\mathrm{d}z}{\mathrm{d}D_L}\right) R\left(\frac{\mathcal{M}}{1+z_\mathrm{rs}}, q, \chi_\mathrm{eff} \,\middle|\, \boldsymbol{\lambda}\right) f(\boldsymbol{p}, D_L).\end{aligned} \tag{3.22}$$

Here, D_c is the comoving distance to the source, and $D_L = (1+z_\mathrm{rs})D_c$ is the luminosity distance. We have assumed a spatially-flat Universe. Following Fishbach and Holz (2017), we have defined $f(\boldsymbol{p}, D_L)$ to be the fraction of events with $\boldsymbol{p}, D_L$ that are detected, averaging over source position (θ, ϕ) and orientation (ι, ψ). We take the detection probability to be one if the expectation value of its SNR exceeds a threshold $\rho_\mathrm{thresh} = 9$ and zero if it does not. In practice, the SNR measured by the detector network will differ from its expectation value due to the noise. The measured SNR would be obtained by maximizing $|z_\mathrm{LIGO}|$ over $t_0, \delta t, \boldsymbol{p}$; its variance over noise realizations is approximately 1 (Allen et al., 2012). The effect that noise fluctuations have on detectability is important only near the boundary of the sensitive volume, and we ignore it for simplicity. The expectation value of the SNR of an event depends on the angles through

$$\rho = A(\theta, \phi, \iota, \psi) \frac{\rho_0(\boldsymbol{p})}{D_{L[\mathrm{Mpc}]}}, \tag{3.23}$$

where $\rho_0(\boldsymbol{p}) = \langle h_0(\boldsymbol{p}) \mid h_0(\boldsymbol{p}) \rangle^{1/2}$ is the SNR that an optimally aligned source at 1 Mpc would have,

and the angular factor $0 \leq A \leq 1$ is given e.g. in Sathyaprakash and Schutz (2009). Then,

$$\begin{aligned} f(\boldsymbol{p}, D_L) &= P(\rho > \rho_{\mathrm{thresh}}) \\ &= P\left(A > \frac{\rho_{\mathrm{thresh}}}{\rho_0(\boldsymbol{p})} D_{L[\mathrm{Mpc}]}\right). \end{aligned} \tag{3.24}$$

We estimate $P(A > A_*)$ from a histogram of A computed over 10^8 isotropically distributed realizations of the angles. In Eq. (3.22), we use the redshift-distance relation given in Adachi and Kasai (2012), taking the values of the cosmological parameters from Ade et al. (2016). We evaluate the D_L integral by quadrature.

To compute $\rho_0(\boldsymbol{p})$, we use a grid over $\boldsymbol{p}$-space and a reference noise PSD. This is valid provided that the shape of the noise curve of the detectors (i.e. the relative values of the PSD at different frequencies) is approximately constant throughout the observation time, at least in the frequency range relevant for BBH detection. We construct the reference PSD as the harmonic mean of the combined-channel PSDs of the first six events: $\sigma^2_{\mathrm{ref}}(f) = \langle \sigma^{-2}_{\mathrm{H},i}(f) + \sigma^{-2}_{\mathrm{L},i}(f) \rangle^{-1}$, where i labels each event and the brackets indicate an average over all the events considered. We consider three BBH events in each observing run, so we expect that this average is representative of the typical PSD during O1 and O2. We find that our results do not sensitively depend on the waveform approximant used.

If one ignored the fact that the source-frame mass depends on redshift, the detector rate would take the form $\Lambda(\boldsymbol{p} \mid \boldsymbol{\lambda}) = R(\boldsymbol{p} \mid \boldsymbol{\lambda}) V(\boldsymbol{p})$, where

$$V(\boldsymbol{p}) = 4\pi \int \mathrm{d}D_L D_L^2 \frac{1}{(1+z_{\mathrm{rs}})^4} \left(1 - \frac{D_L}{1+z} \frac{\mathrm{d}z}{\mathrm{d}D_L}\right) f(\boldsymbol{p}, D_L) \tag{3.25}$$

is the sensitive (comoving-) volume of the detector network. Fig. 3.3 shows the sensitive volume computed with the `SEOBNRv4_ROM` approximant. V indeed depends on all the three parameters $\mathcal{M}, q, \chi_{\mathrm{eff}}$. As already pointed out in Fishbach and Holz (2017), the mass dependence follows an approximate power law $V \propto \mathcal{M}^{2.2}$ for $q \gtrsim 0.5$, i.e. heavier BBH mergers are louder. Moreover, events with large χ_{eff} (where the spins are aligned with the orbital angular momentum) are also louder, because χ_{eff} first enters the post-Newtonian expansion as an effective force that is repulsive for $\chi_{\mathrm{eff}} > 0$ (Campanelli et al., 2006). This effect is irrelevant for the dynamics of the inspiral while the BHs are far apart, but it means that the signal lasts longer in the detector band before the plunge, and thus more SNR is accumulated. Finally, the dependence with q at fixed $\mathcal{M}$ is very weak if $q \gtrsim 0.5$, but the sensitive volume drops rather strongly for smaller q.

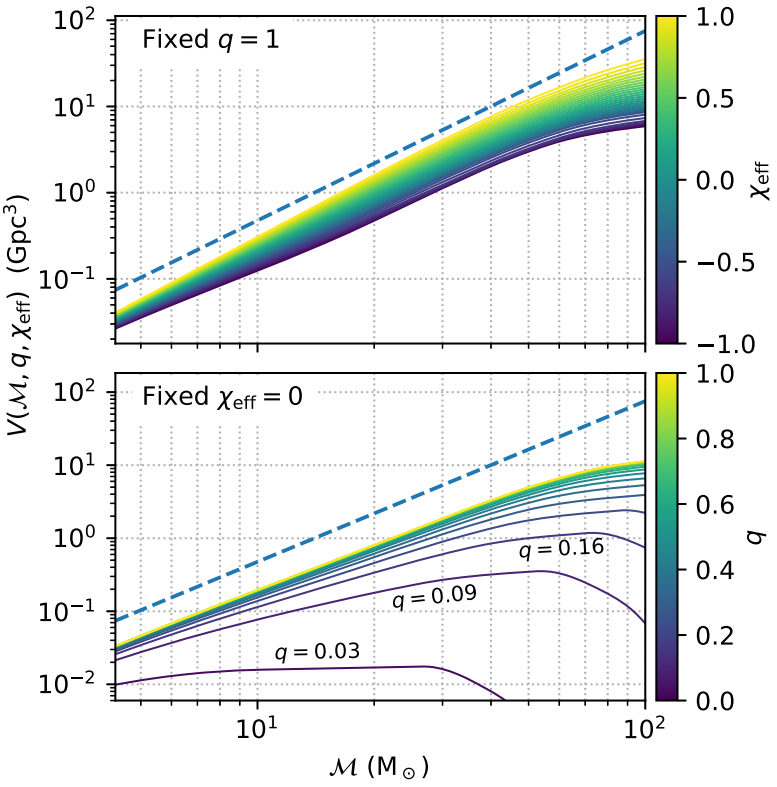


Figure 3.3: Dependence of the detector network sensitive volume on the parameters $\boldsymbol{p} = (\mathcal{M}, q, \chi_{\rm eff})$. The solid lines show $V(\mathcal{M})$ for several values of $\chi_{\rm eff}$ on a $q = 1$ slice (top), or several values of q on a slice of constant $\chi_{\rm eff} = 0$ (bottom). The dashed line shows a power law $\propto \mathcal{M}^{2.2}$ dependence for reference.

Another possible source of bias in the inferred merger rates would arise if the effectualness of the template bank in recovering signals depended sensitively on the parameters. Although this effect is present, its magnitude is much smaller than the sensitive volume dependence for BBHs with $\mathcal{M} < 100\,\mathrm{M}_\odot$ computed here (Canton and Harry, 2017).

3.3.2 Model likelihood

The likelihood that a specific model for the merger rates will have the observed set of detections as an outcome is

$$P(\{d_i\} \mid \boldsymbol{\lambda}) \propto \prod_{i \in \text{events}} \int \mathrm{d}\boldsymbol{p} P(d_i \mid \boldsymbol{p}, \mathcal{D}) P(\boldsymbol{p} \mid \mathcal{D}, \boldsymbol{\lambda}). \tag{3.26}$$

Here, $P(d_i \mid \boldsymbol{p}, \mathcal{D})$ is the single-event likelihood $\overline{\mathcal{L}}_i$ given by (3.21), and $P(\boldsymbol{p} \mid \mathcal{D}, \boldsymbol{\lambda})$ is the detection prior for the event parameters $\boldsymbol{p}$ according to the model $\boldsymbol{\lambda}$, which is proportional to the detector

rate (3.22) but normalized to $\int \mathrm{d}\boldsymbol{p} P(\boldsymbol{p} \mid \mathcal{D}, \boldsymbol{\lambda}) = 1$. Using this we can rewrite (3.26) as

$$P(\{d_i\} \mid \boldsymbol{\lambda}) \propto \frac{\prod_i \int \mathrm{d}\boldsymbol{p} \overline{\mathcal{L}}_i(\boldsymbol{p}) \Lambda(\boldsymbol{p} \mid \boldsymbol{\lambda})}{\left[\int \mathrm{d}\boldsymbol{p} \Lambda(\boldsymbol{p} \mid \boldsymbol{\lambda})\right]^{N_{\mathrm{events}}}}, \tag{3.27}$$

which we can compute on a grid in $\boldsymbol{\lambda}$-space, given a merger-rate model $R(\boldsymbol{p} \mid \boldsymbol{\lambda})$. The $\boldsymbol{p}$ integrals in the numerator of (3.27) only have support near the measured parameter values of each event, since $\overline{\mathcal{L}}_i(\boldsymbol{p})$ is suppressed elsewhere. The integral in the denominator runs over all sensitive parameter space, which we take to be $\mathcal{M} \in [4.3\,\mathrm{M}_\odot, 100\,\mathrm{M}_\odot], q \in [0.03, 1], \chi_{\mathrm{eff}} \in [-1, 1]$. The lower limit on $\mathcal{M}$ holds if all astrophysical BHs have $m > 5\,\mathrm{M}_\odot$. We chose the lower bound on q to include the regions where the reported events have support. Note that such low values of q are outside the calibration region of the approximants (Khan et al., 2016; Bohé et al., 2017), although the effect for the purpose of SNR estimation should be minor, and also the approximation that $(l, m) = (2, \pm 2)$ is less accurate in that regime. In any case, those mass ratios are highly suppressed in the sensitive-volume weighting (Eq. (3.25), Fig. 3.3).

To get a posterior distribution for the $\boldsymbol{\lambda}$-parameters, one should multiply the likelihood by a prior $P_{\mathrm{prior}}(\boldsymbol{\lambda})$. We will take those priors to be flat.

We emphasize again that the model likelihood (3.27) accounts both for selection effects due to the sensitive volume and for the fact that the priors that one has to use to estimate the parameters of each event depend on the merger-rate model that one is considering.

3.4 Astrophysical implications

With only 10 detections made so far, it is not yet feasible to constrain population models that are too complicated. For this reason, we consider several simple models that intend to probe the different variables separately, and apply Eq. (3.27) to put constraints on their parameters (see Taylor and Gerosa (2018) for an alternative framework tailored to constraining detailed models).

As before, we repeated all analyses using two different waveform approximants (`SEOBNRv4_ROM` or `IMRPhenomD`). We find that our results are robust to these choices, so below we only show the results for the `SEOBNRv4_ROM` approximant.

3.4.1 Spin distribution

The distribution of spins of the merger events is currently one of the more informative data the LVC has presented. First, the spin distribution might allow us to distinguish between various formation

channels. For example in scenarios where black holes are dynamically captured into binaries one expects each spin to be randomly oriented. For field binaries spins might tend to be aligned with the orbital angular momentum. Tides in binary systems before the second black hole forms might spin up the secondary and align it with the orbital angular momentum. For a chemically homogeneous evolution of the stars to happen, high spins are required. Thus the LVC measurements of $\chi_{\rm eff}$ can potentially provide very interesting constraints.

Second, one could try to ascertain whether the properties of the merging black holes are similar to those of black holes in X-ray binaries. In particular there is some indication that local black holes are rotating fast. The middle panel of Fig. 3.1 shows the constraints on mass and spin of a collection of black holes in XRBs from X-ray measurements. Heavy black holes in persistent sources, i.e. with heavy companions, which are the natural progenitors of the LIGO/Virgo sources, are close to maximally spinning. Furthermore this spin is usually interpreted as being natal and thus perhaps should apply to the secondary black hole as well. By comparison, the $\chi_{\rm eff}$ reported by the LVC seem rather low. Of course $\chi_{\rm eff}$ constrains only one of the components of the spin and combines both black holes with weights depending on the mass ratio. We will try to use the likelihoods we have computed for the LIGO events to say something about the spin magnitudes and orientations assuming they all come from the same population.

Gaussian $\chi_{\rm eff}$ rate model

In order to understand what the data are already telling us about the distribution of spins we first consider a merger-rate model which is simply a (truncated) Gaussian in $\chi_{\rm eff}$,

$$R(\chi_{\rm eff} \mid \overline{\chi_{\rm eff}}, \sigma_{\chi_{\rm eff}}) \propto G(\chi_{\rm eff} - \overline{\chi_{\rm eff}}, \sigma_{\chi_{\rm eff}}), \qquad |\chi_{\rm eff}| < 1; \tag{3.28}$$

we will use $G(x, \sigma)$ to note the Gaussian distribution

$$G(x, \sigma) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

We allow a nonzero mean, as expected for example from an isolated-binary formation scenario, and a dispersion $\sigma_{\chi_{\rm eff}}$ whose value can help us constrain the typical magnitude of the individual spins. The relevant values of $\sigma_{\chi_{\rm eff}}$ turn out to be $\lesssim 0.2$, so in the following we will make no distinction between $\sigma^2_{\chi_{\rm eff}}$ and the variance of the truncated Gaussian. For simplicity, in this example we adopt a uniform prior in $\mathcal{M}, q$.

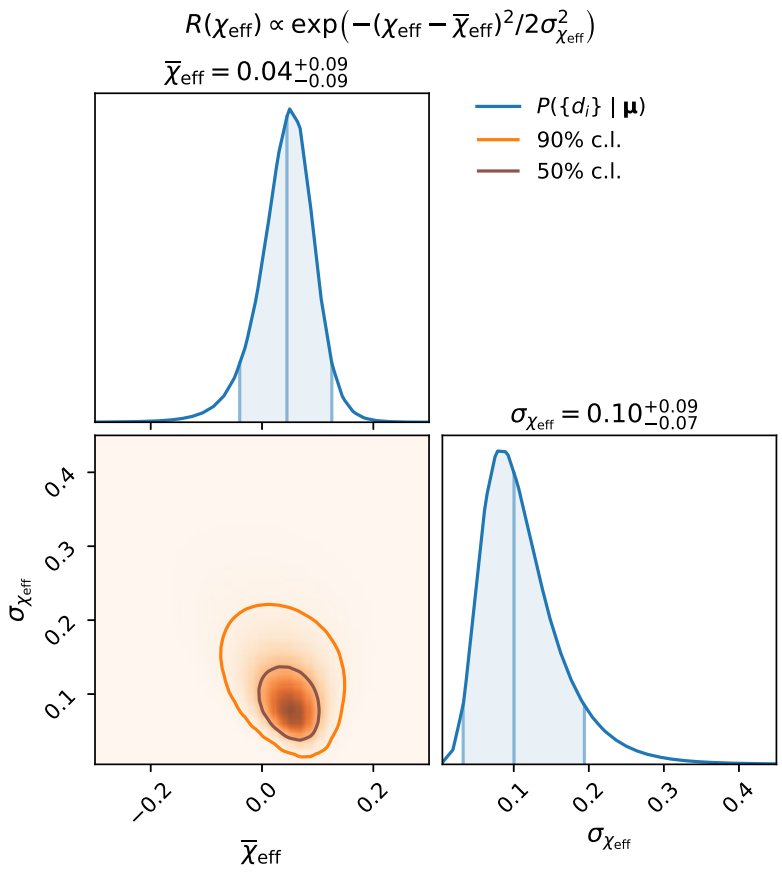


Figure 3.4: Merger-rate model parameter likelihood for a Gaussian distribution of $\chi_{\rm eff}$ given by (3.28). Probability contours enclosing 50% and 90% of the distribution are shown in the two-dimensional plot. The one-dimensional plots show the single-parameter marginalized likelihood. The vertical lines show the marginalized distribution median and the minimal 90% probability interval.

The $\boldsymbol{\lambda}$-parameter likelihood is shown in Fig. 3.4. The distribution is consistent with having zero mean, with a mild preference for positive values. The figure also shows that $\sigma_{\chi_{\rm eff}} = 0$ is inconsistent with the data. We find an upper 90% bound $\sigma_{\chi_{\rm eff}} < 0.19$.

To interpret these results in terms of the distributions of the individual spins we assume that each spin is drawn from a distribution with average spin $\overline{a}$ and dispersion σ_a, with an angle relative to the angular momentum whose cosine has a mean $\overline{\mu}$ and a dispersion σ_μ. We will allow the two angles to potentially be correlated so that $\langle \mu_1 \mu_2 \rangle = r_\mu \sigma_\mu^2$. In this case we can compute the mean and variance of $\chi_{\rm eff}$:

$$\begin{aligned} \overline{\chi_{\rm eff}} &= \overline{a}\,\overline{\mu} \\ \sigma^2_{\chi_{\rm eff}} &= \frac{1+q^2}{(1+q)^2}\overline{\mu}^2\sigma_a^2 + \left(\frac{1+q^2+2qr_\mu}{(1+q)^2}\overline{a}^2 + \frac{1+q^2}{(1+q)^2}\sigma_a^2\right)\sigma_\mu^2. \end{aligned} \tag{3.29}$$

We can first consider situations in which there is no preference for aligned spins, $\overline{\mu} = \overline{\chi_{\rm eff}} = 0$ and $\sigma_\mu^2 = 1/3$, such as BBH populations that would arise in dynamical capture scenarios. The variance of $\chi_{\rm eff}$ still depends on r_μ. In the two limits of no correlation ($r_\mu = 0$) and perfect correlation ($r_\mu = 1$) we get:

$$\begin{aligned}\sigma^2_{\chi_{\rm eff}}(r_\mu = 0) &= \frac{1}{3}\left(\overline{a}^2 + \sigma_a^2\right)\frac{1+q^2}{(1+q)^2},\\ \sigma^2_{\chi_{\rm eff}}(r_\mu = 1) &= \frac{1}{3}\left(\overline{a}^2 + \sigma_a^2\frac{1+q^2}{(1+q)^2}\right)\end{aligned} \tag{3.30}$$

The dynamical capture scenario would correspond to $r_\mu = 0$. One can envision a situation for a binary star progenitor in which the spins are misaligned with the orbit but have similar directions, leading to $r_\mu = 1$. For example, natal kicks at BH formation if the supernova explosion is asymmetric, or perhaps the tidal interaction with a third body (Rodriguez and Antonini, 2018), would lead to a spin–orbit misalignment. Spins that are misaligned with the orbit precess, which would spoil the spin–spin alignment. Even so, precession conserves $\chi_{\rm eff}$ to a large extent (Apostolatos et al., 1994; Racine, 2008) so it does not affect the $\chi_{\rm eff}$ distribution. This holds as long as the orbital angular momentum is bigger than the BH spins, which is generally the case unless the mass ratio is very small, $q \lesssim \chi_1 v/c$ (so it is valid for the sources LIGO/Virgo are most sensitive to, see Fig. 3.3). The limit of completely random misalignment between spin and orbital angular momentum considered here is extreme, for example the necessary kicks would unbind most systems. In any event we consider it to understand the limits of the constraints we get.

We first consider whether the data allow for high spins as might be hinted by XRBs. For equal mass ratios and assuming $\overline{a} \approx 1$, these dispersions would be $\sigma_{\chi_{\rm eff}}(r_\mu = 0, q = 1) = 1/\sqrt{6} \approx 0.4$ and $\sigma_{\chi_{\rm eff}}(r_\mu = 1, q = 1) = 1/\sqrt{3} \approx 0.6$. Both values appear to be too large, indicating that the data already do not favour large spins even if both of them are randomly oriented. One could consider small mass-ratios, in which case for high spins the dispersions become $\sigma_{\chi_{\rm eff}}(r_\mu = 0, q \ll 1) \sim \sigma_{\chi_{\rm eff}}(r_\mu = 1, q \ll 1) \sim 1/\sqrt{3} \approx 0.6$, which again seems disfavoured.

In the binary progenitor scenario it is believed that the BHs have a preference for being aligned with the orbital angular momentum. As an extreme version let us consider the case of perfect spin–orbit alignment $\overline{\mu} = 1, \sigma_\mu = 0$. The mean and variance become:

$$\begin{aligned}\overline{\chi_{\rm eff}} &= \overline{a}\\ \sigma^2_{\chi_{\rm eff}} &= \frac{1+q^2}{(1+q)^2}\sigma_a^2.\end{aligned} \tag{3.31}$$

Fig. 3.4 shows that the data are consistent with having zero mean for $\chi_{\rm eff}$, and at most around $\overline{\chi_{\rm eff}} \lesssim 0.1$, thus $\overline{a} \lesssim 0.1$. The data also require a non-zero variance, this demands $\sigma_a \neq 0$ (because we have assumed perfect alignment, $\sigma_\mu = 0$ and the only source of variance for $\chi_{\rm eff}$ in Eq. (3.29) remains σ_a). Since $a > 0$, in this limit of small spins one cannot have $\sigma_a \gg a$. Thus, from Fig. 3.4 the only viable region for the aligned scenario is $\sigma_a \sim \overline{a} \sim 0.1$

Rate model for individual spins

Although we have already obtained most of the interesting physical conclusions from the previous simple analysis perhaps it is warranted to be a bit more careful with the dependence on q. Both the model predictions Eq. (3.29) and the inferred $\chi_{\rm eff}$ for the events (Fig. 3.1) depend on q. In the case of the data, for the lower-mass events there is a very strong degeneracy between $\chi_{\rm eff}$ and q. It is only for small mass-ratios that the spin parameter becomes large. Therefore, when we look at the constraints on $\sigma_{\chi_{\rm eff}}$ larger values are allowed due to the possibility that the events have small q. But the models tend to predict a larger variance in this regime, because in this limit only one of the two spins contributes to $\chi_{\rm eff}$ and there is no possibility of cancellation. Thus, it is worth considering directly a prior on the individual spins rather than on $\chi_{\rm eff}$ so that the q dependence is automatically incorporated.

With this small number of events we do not want to consider very complicated rate models. We restrict ourselves to a two-parameter model that explores at the same time the typical value of the individual spin magnitudes as well as their alignment with the orbit. We will consider a rate model where each spin–orbit alignment μ is uniformly distributed between $\mu_{\rm min}$ and 1. For the spin magnitude we know that in the limit of perfect alignment the data require that the spin distribution have some variance. Thus we will consider a uniform distribution between $a_* - \Delta_a$ and $a_* + \Delta_a$ and move a_*. When a_* approaches the boundaries only $0 < a < 1$ is allowed. We will take $\Delta_a = 0.1$.

The likelihood for $a_*, \overline{\mu}$ is shown in Fig. 3.5. We find that the typical spin magnitudes have a preferred value $a_* \approx 0.2 \pm 0.2$, in agreement with Wysocki et al. (2019). The alignment has an upper 90% bound of $\overline{\mu} \lesssim 0.6$, but there is support all the way up to $\overline{\mu} = 1$ if the typical spins are $a_* \approx 0.1$, consistent with our estimate in §3.4.1.

Tidally-locked progenitor

We now consider the effects of tides in the binary. For a field binary, before becoming a BBH the progenitor was a binary star where the companion of the primary BH probably was a Wolf–Rayet star (the core of a star that lost its envelope). The minimum initial distance required for the BBH to

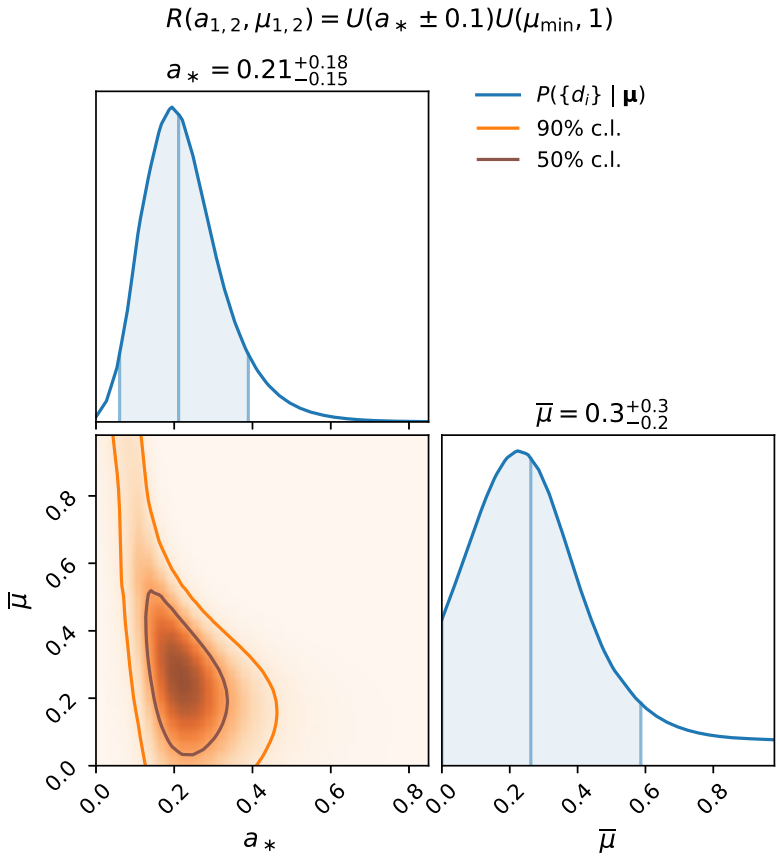


Figure 3.5: Parameter likelihood for a merger-rate model in which the BBHs have spin magnitude a uniformly distributed between $a_* - 0.1$ and $a_* + 0.1$, and alignment between the BH spin and the orbital angular momentum uniformly distributed between $\mu_{\min} = 1 - 2\overline{\mu}$ and 1.

merge in the age of the Universe due to the emission of gravitational radiation is comparable to the minimum distance required to tidally lock the companion star within its lifetime, and the time-scales for these two processes are very steep functions of the distance, so two distinct subpopulations are expected (Kushnir et al., 2016; Hotokezaka and Piran, 2017; Zaldarriaga et al., 2017; Qin et al., 2018). In systems where the merger time is shorter than $\sim 10^8$ years, the secondary would be tidally locked and rapidly spinning. If the merger time is longer, tides are too weak. We might thus expect two distinct populations.

A strong natal kick to the second BH could misalign the orbit, but we ignore this case as it is rather unlikely given that at the time of the second explosion the binary is already tight, and thus the needed kick velocity to produce large misalignments is too large. We assign the high spin to the secondary (lighter) BH.

In Figs. 3.1 and 3.B.1, 3.B.2 we used the prescription $\chi_1 = 0, \chi_2 = 1$ as a proxy for the case in which the Wolf–Rayet star is successfully locked, which defines the curve $\chi_{\rm eff} = q/(1+q)$ shown. We find that GW170729, GW151226, GW151012 and GW170608 are consistent with this proxy. The latter three are consistent if their mass ratio is low, while all the other, heavier events are more consistent with $q \sim 1$. Thus this would be slightly unlikely if all the events belong to the same population. With the current error bars and small number of events this is at most a qualitative hint. To further illustrate this, in Fig. 3.6 we show the likelihood as a function of a rescaled $\chi_{\rm eff}$ in units of the proxy $q/(1+q)$. The rescaled effective spin can take values between $\pm(1+q)/q$, which for a reference $q = 1/2$ is ± 3. Zaldarriaga et al. (2017) predicted a bimodal distribution in this variable, with peaks at 0 and 1 corresponding to the two subpopulations.

We implement a two-parameter model of this scenario, described by the fraction f of systems where the progenitor has been tidally-locked, and the standard deviation $\sigma_{\chi_{\rm eff}}$ in the effective spin of the other subpopulation. We assume that the aligned spin of each black hole either comes from the same zero-mean Gaussian distribution, or is unity for a fraction f of the secondaries. That is, the standard deviation of the distributions of individual spin and effective spin are related by

$$\sigma_{\chi_1} = \sigma_{\chi_2} = \frac{1+q}{\sqrt{1+q^2}}\sigma_{\chi_{\rm eff}}, \tag{3.32}$$

valid for the black holes that do not come from a tidally-locked progenitor. This results in a superposition of two Gaussian distributions for $\chi_{\rm eff}$:

$$R(\chi_{\rm eff}, q \mid f, \sigma_{\chi_{\rm eff}}) = (1-f)G(\chi_{\rm eff}, \sigma_{\chi_{\rm eff}}) + fG\left(\chi_{\rm eff} - \frac{q}{1+q}, \frac{\sigma_{\chi_{\rm eff}}}{\sqrt{1+q^2}}\right). \tag{3.33}$$

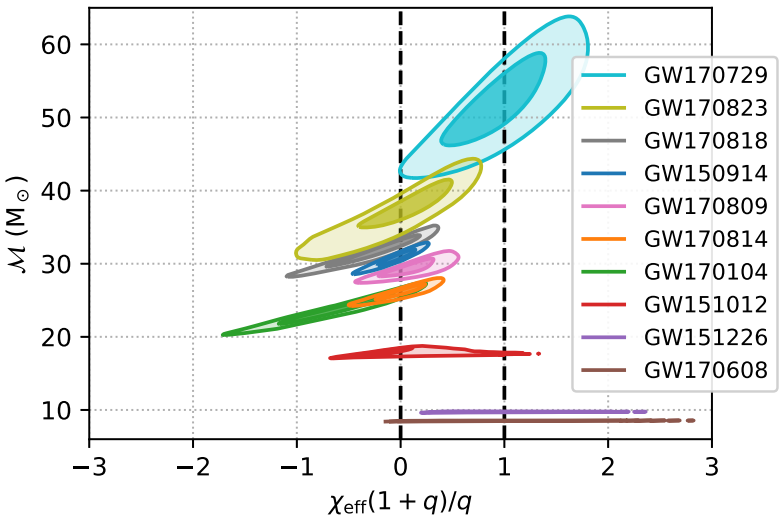


Figure 3.6: Single-event likelihood in terms of $\chi_{\rm eff}$ rescaled by the tidally-locked-progenitor proxy $q/(1+q)$, and marginalized over q. This channel predicts a bimodal distribution in this variable, with peaks at 0 and 1.

The second term describes the locked-progenitor subpopulation, it has a positive mean due to the maximally-spinning secondary, and a smaller variance because in this case only the primary is random.

It should be noted that varying the prescriptions used for stellar winds and tides it can also be possible to obtain intermediate distributions where the secondary may not be maximally spinning after being tidally locked (Qin et al., 2018). However, it would be hard to constrain models more complicated than (3.33) with the present number of events; and the upper bound on f would get stronger as long as we interpret f as the fraction of maximally-spinning secondaries.

The likelihood for $f, \sigma_{\chi_{\rm eff}}$ is shown in Fig. 3.7. We see that the data are consistent with $f = 0$ and puts an upper 90% bound $f < 0.3$. One can see a hint of two peaks, one in which the random spin is used to explain all the events, leading to a larger variance, and the other where the higher-spin events are explained using tides and the random component of the spins is low. The current number of events is too small to discriminate but it should become possible in the next LIGO run. Furthermore if a negative $\chi_{\rm eff}$ were to be observed, especially a large one compared with the width of the random component, then this scenario would be disfavoured.

Trends with mass and GW170729

It is important to notice that in the analysis so far we have assumed that all the LIGO events are samples from a unique population. Fig. 3.1 shows some noticeable differences between the light and

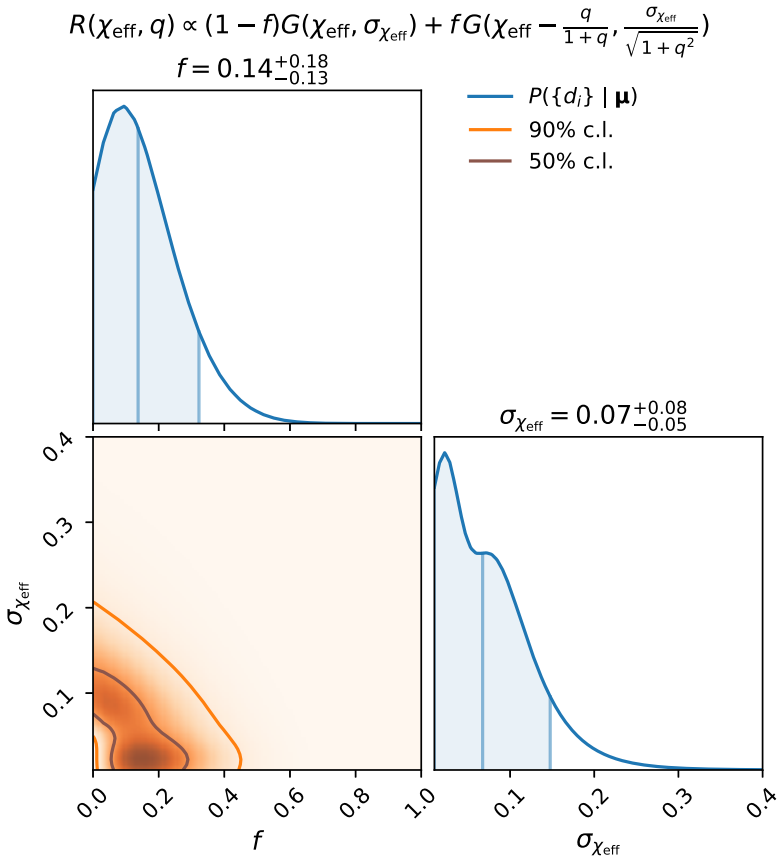


Figure 3.7: Parameter likelihood for a merger-rate model where a fraction f of the sources comes from a tidally-locked progenitor. The distribution of $\chi_{\rm eff}$ is given by (3.33).

heavy systems. The most obvious difference is just related to the parameter combination than can be measured best, the change in the degeneracy direction from being between spin and mass ratio to being between spin and mass. Of course this is just a feature of the way the strain depends on the parameters.

However, prior to the detection of GW170729, a hint that the lightest systems might be the ones with a clearer positive spin was apparent. With such small number of events we did not dare start splitting the sample in subsets and especially in such an a posteriori fashion. However, a trend with mass could be expected on theoretical grounds and is sometimes used as input in population synthesis models (see for example figure 1 of Belczynski et al. (2017)). Of course one has to worry about those inputs because even though they might be physically motivated, some of them were chosen after seeing the first set of LIGO BBHs.

GW170729 is heavy and has a moderately high spin, which breaks this potential trend and could mean either that there is no such trend or that a different physical mechanism originated its spin. Formation scenarios that would naturally account for this event include tidal-locking after a common-envelope phase as discussed in §3.4.1; a chemically-homogeneous formation, which predicts masses in the range of GW170729, near-equal mass ratio, aligned spins and a peak merger rate at redshift 0.5 (Mandel and de Mink, 2016); or a repeated merger within the globular cluster scenario, which would also have high mass and spin (although randomly aligned) and would represent a fraction of the mergers coming from globular clusters that could be as high as 20% if the spins at birth are small (Rodriguez et al., 2018).

3.4.2 Mass distribution

Another interesting question is the distribution in mass and the potential lack of heavy systems. This was already pointed out in Fishbach and Holz (2017). Here we repeat that analysis with six additional events, allowing for spin (through $\chi_{\rm eff}$ only) and including the source-frame mass dependence on redshift. For simplicity we directly model the distribution as a function of $\mathcal{M}$ and adopt a power-law prior with a cutoff. One could model the distribution of the individual masses and make further assumptions about how correlated the two masses are. We feel that this is unnecessary at this stage as the events have mass ratios $q \sim 1$ and a cutoff in the mass distribution would also lead to a cutoff in the chirp mass distribution at a very similar mass, especially given the current errors and small number of events. We adopt a $5\,\mathrm{M}_\odot$ low-mass cutoff and a free high-mass cutoff parameter $\mathcal{M}_{\rm max}$ for the model distribution. We recall that we assumed that the detector network is sensitive

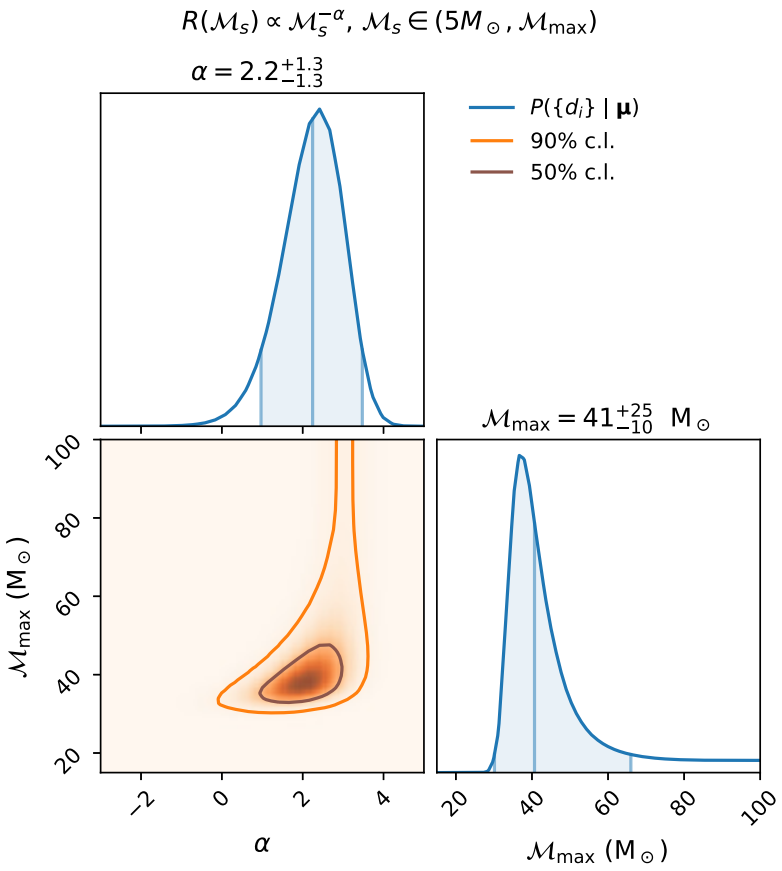


Figure 3.8: Parameter likelihood for a merger-rate model where the source-frame chirp-mass distribution follows a power law $\mathcal{M}_s^{-\alpha}$, with a low-mass cutoff at $5\,\mathrm{M}_\odot$ and a high-mass cutoff at $\mathcal{M}_{\rm max}$.

to events with $\mathcal{M} < 100\,\mathrm{M}_\odot$, which is similar to and less conservative than the $m_1 + m_2 < 200\,\mathrm{M}_\odot$ case analysed in Fishbach and Holz (2017).

We show our results in Fig. 3.8, which qualitatively agree with Fishbach and Holz (2017) and give a preliminary indication of the presence of a maximum mass. The distribution median is $\mathcal{M}_{\rm max} = 41\,\mathrm{M}_\odot$, which would correspond to $m_{1,2} = 47\,\mathrm{M}_\odot$ if $q = 1$, also in agreement with Abbott et al. (2019b). The lower bound on $\mathcal{M}_{\rm max}$ is determined by the heaviest event detected, GW170729.

One potential caveat is that the glitch background in LIGO might increase with mass, partially compensating the increase in volume up to which heavy systems can be seen. To investigate this we estimate how sensitive our results are to an increase in the background. From fig. 9 of Abbott et al. (2016c), the background of triggers during O1 is a steeply decreasing function of the detector network SNR ρ and it is approximated by a straight line for $\log_{10} N_{\rm bg}(\rho)$. This also holds for unmodeled searches (Lynch et al., 2018). The background is dominated by detector glitches. GW signals from

heavy BBHs have short duration and fewer cycles in the detector band, so glitches can more easily resemble them. The background level is then dependent on the mass scale and a stricter threshold on the SNR might be necessary for heavier events, which has the effect of reducing the sensitive volume relative to our previous estimate (3.25). We can make a simple estimate of this effect as follows. In an Euclidean spacetime (setting $z_{\rm rs}(D_L) = 0$ in Eq. (3.25)), the sensitive volume scales as $V(\boldsymbol{p}) \propto \rho_0^3(\boldsymbol{p})$, since the SNR of a source decays as D^{-1}. If we allow for a parameter-dependent threshold on the SNR, V gets a correction

$$\frac{V(\boldsymbol{p})}{V(\boldsymbol{p_0})} = \left(\frac{\rho_0(\boldsymbol{p})}{\rho_0(\boldsymbol{p_0})}\right)^3 \left(\frac{\rho_{\rm thresh}(\boldsymbol{p_0})}{\rho_{\rm thresh}(\boldsymbol{p})}\right)^3, \tag{3.34}$$

where $\boldsymbol{p_0}$ is some reference parameter value. The first term in the right hand side comes from Eq. (3.25), and the second describes the reduction in sensitive volume if the threshold $\rho_{\rm thresh}(\boldsymbol{p}) > \rho_{\rm thresh}(\boldsymbol{p_0})$. In an expanding Universe, the $(1 + z_{\rm rs})^{-4}$ term in Eq. (3.25) suppresses the large-luminosity-distances contribution to V, so the relative decrease in sensitive volume due to raising the threshold for large masses is smaller than the estimate (3.34).

As an example, if the glitch rate increased from the one reported in fig. 9 of Abbott et al. (2016c) by a factor of 100 above a certain chirp mass, the SNR threshold would have to be raised by $\Delta\rho_{\rm thresh} \approx 1.07$ for those events to have the same false-alarm rate. Assuming a detection threshold of $\rho_{\rm thresh}(\boldsymbol{p_0}) = 9$, the sensitive-volume correction factor in Eq. (3.34) is 0.71. There is no indication of such a dramatic increase in the background even when going beyond $\mathcal{M} > 100\,\mathrm{M}_\odot$ (Abbott et al., 2017a). A thorough test with an injection campaign was performed by Abbott et al. (2019b) and found correction factors consistent with this estimate. Thus we believe the hint of a cutoff mass to be a robust result.

3.4.3 Mass-ratio distribution

Finally, we consider what we can say about the mass-ratio distribution. The likelihoods themselves are rather flat in q so we do not expect particularly good constraints. We consider a power law in q and try to constrain the exponent. We present our results in Fig. 3.9. We find that typical mass ratios below 0.7 seem disfavoured. For comparison, the average mass ratio of a distribution where the binary masses are independently taken from a power-law $P(m) \propto m^{-\alpha}$ is $\overline{q} = 1 - 1/\alpha$. From Fig. 3.8 we obtain that $\alpha \approx 3$ is favoured by the data if we impose no cutoff ($\mathcal{M}_{\rm max} \to \infty$); independent draws from this distribution which would yield $\overline{q} = 2/3$, consistent with our lower bound on $\overline{q}$. If we use the Salpeter mass function $\alpha = 2.35$, then $\overline{q} = 0.575$, which is disfavoured.

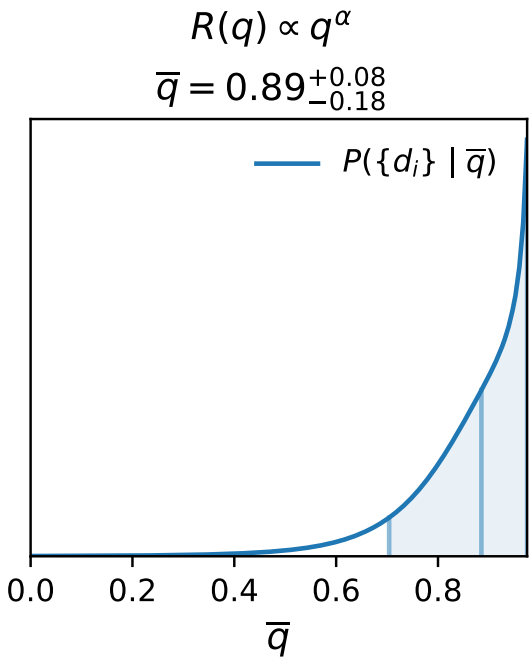


Figure 3.9: Parameter likelihood for the average mass-ratio $\overline{q}$, in a merger-rate model where the mass-ratio distribution follows a power law q^{α}, with α adjusted to yield $\langle q \rangle = \overline{q}$.

A uniform distribution in q would have $\overline{q} = 1/2$, also disfavoured. We note that the statement that equal mass-ratios are favoured holds for the physical merger-rate in the Universe, even after accounting for selection effects due to the higher sensitive volume of the detector to those systems.

Due to the degeneracy between spin and mass ratio we could consider the combination of a Gaussian in $\chi_{\rm eff}$ and a power law in q. However, when this is done we find the constraints on the variance of the Gaussian and the mean $\overline{q}$ to be the same as those we report in Figs. 3.4 and 3.9.

3.5 Discussion

It is clear that even with the small number of events reported so far the LIGO–Virgo data already provide hints that have interesting astrophysical consequences. Most interestingly, the data suggest that the BHs do not all spin rapidly, perhaps in tension with the simplest interpretation of the X-ray binary results. This is even true in dynamical scenarios where the spins are randomly oriented and even more so in the field binary case where there might be a tendency towards alignment between spins and angular momentum.

These results are largely consistent with those presented in Farr et al. (2017), where aligned high spin distributions were disfavoured. Quantitatively we differ because we have more events, have a

better estimate of the likelihoods for each event including the degeneracy between mass ratio and spin, and rather than comparing discrete models we have continuous parameters that connect them, allowing us to perhaps have a better sense of the typical spins that are favoured or disfavoured in individual scenarios.

There are many potential explanations for the difference between the typical spins of high-mass XRBs and the LIGO/Virgo systems. Of course both data sets contain a small number of events and thus this could be a statistical fluke. Perhaps there is an unrecognized systematic in one or both measurements. There could be astrophysical explanations, perhaps these systems come from different populations. There could even be exotic explanations such as the effect of an axion-type particle that through superradiance extracts energy from rotating black holes to produce a cloud of axions around them.

At the current time the simple test of seeing if the spin distribution has a tendency towards positive spins is not powerful enough. We do not detect a mean to the distribution and thus cannot use that to distinguish between field binary and dynamical scenarios. Regarding tides, four of the events are consistent with having a maximally spinning secondary, although for the lighter three this only happens if the mass ratio for these events is lower than for the heavy ones, and in particular GW151012 is only marginally consistent. With the data at hand we cannot make a conclusive statement but again it is clear that the next few events are going to be very interesting.

Regarding the mass distribution we confirm previous findings that there is a hint of a cutoff in mass. The distribution of mass ratios is not so well constrained, in part due to the degeneracy with the spin, but distributions where the BBHs have comparable masses are favoured.

The code and data used in this work are publicly available at `https://github.com/jroulet/constraints_bbh_populations`.

Appendix

3.A Selection bias

In this appendix we discuss in further detail the equivalence between our derivation of Eq. (3.27) and others present in the literature. Letting $\mathcal{D}$ be the proposition that the event was detected (and $\neg\mathcal{D}$ its negation), we can write the marginalized likelihood as

$$\begin{aligned} P(d \mid \boldsymbol{p}) &= P(d \mid \mathcal{D}, \boldsymbol{p}) P(\mathcal{D} \mid \boldsymbol{p}) + P(d \mid \neg\mathcal{D}, \boldsymbol{p}) P(\neg\mathcal{D} \mid \boldsymbol{p}) \\ &= P(d \mid \mathcal{D}, \boldsymbol{p}) P(\mathcal{D} \mid \boldsymbol{p}) + P(\neg\mathcal{D} \mid d, \boldsymbol{p}) P(d \mid \boldsymbol{p}) \\ &= P(d \mid \mathcal{D}, \boldsymbol{p}) P(\mathcal{D} \mid \boldsymbol{p}). \end{aligned} \tag{3.35}$$

The last equality follows because the data segments we analysed were detections, and the criterion for detection depends only on the data, so $P(\neg\mathcal{D} \mid d) = 0$. The term $P(\mathcal{D} \mid \boldsymbol{p})$ is the observational bias and can be computed as

$$\begin{aligned} P(\mathcal{D} \mid \boldsymbol{p}) &= \int_{a_{0,\min}}^{\infty} P(a_0 \mid \boldsymbol{p}) \mathrm{d}a_0 \\ &= \int_0^{D_h(\boldsymbol{p})} P(D) \mathrm{d}D \\ &\propto V(\boldsymbol{p}), \end{aligned} \tag{3.36}$$

where $D_h(\boldsymbol{p}) = \sqrt{\langle h_0 \mid h_0 \rangle(\boldsymbol{p})}\,\mathrm{Mpc}/a_{0,\min}$. The probability of detection depends on the parameters because we are keeping only a subset $\boldsymbol{p}$ of the parameters, and in particular marginalizing over distance and angles (cf. Loredo (2004); Mandel et al. (2019), keeping all the parameters would lead to a detection probability of 1 so they can omit this term). Using Eqs. (3.35) and (3.36) we can readily show the equivalence between our Eq. (3.27) and eq. (7) of Mandel et al. (2019). The

expressions differ only in the integrands of the numerators of each term in the product:

$$\begin{aligned} P(d_i \mid \boldsymbol{p}, \mathcal{D}_i)\Lambda(\boldsymbol{p} \mid \boldsymbol{\lambda}) &= P(d_i \mid \boldsymbol{p}, \mathcal{D}_i)V(\boldsymbol{p})R(\boldsymbol{p} \mid \boldsymbol{\lambda}) \\ &\propto P(d_i \mid \boldsymbol{p}, \mathcal{D}_i)P(\mathcal{D} \mid \boldsymbol{p})R(\boldsymbol{p} \mid \boldsymbol{\lambda}) \\ &= P(d_i \mid \boldsymbol{p})R(\boldsymbol{p} \mid \boldsymbol{\lambda}) \end{aligned} \tag{3.37}$$

which is the form found in Mandel et al. (2019). Note that Eq. (3.37) is the posterior for $\boldsymbol{p}$ under a prior labeled $\boldsymbol{\lambda}$. In Fig. 3.B.1 we use the LIGO prior as an example.

We have expressed Eq. (3.27) in terms only of $P(d \mid \boldsymbol{p}, \mathcal{D})$ and $\Lambda(\boldsymbol{p} \mid \boldsymbol{\lambda})$ instead of the more physically-interesting quantities $P(d \mid \boldsymbol{p})$ and $R(\boldsymbol{p} \mid \boldsymbol{\lambda})$. Our motivation is that this expression is perhaps more natural from the perspective of the observers, since the outcome of the observations can only depend on the event rate at the detector and the events analyzed will necessarily be conditioned to detection.

3.B Comparison with LVC results

In Fig. 3.B.1 we show the likelihood on the $\mathcal{M}, q, \chi_{\rm eff}$ parameter space for each event, and the posterior distribution computed using the same prior as LIGO to facilitate the comparison. Our reanalysis recovers both the one- and two-dimensional marginalized posteriors accurately, which is compelling evidence that our approximations are working well and capture the degeneracies present.

The prior LIGO used for the astrophysical distribution of parameters is uniform in the individual masses, spin magnitudes and cosine of spin tilts. This induces a nonuniform prior on the variables we adopted. The transformation from m_1, m_2 to $\mathcal{M}, q$ is given by Eq. (3.2), taking its Jacobian yields

$$P_{\rm LIGO\,prior}(\mathcal{M}, q) \propto \frac{1}{q}\left[\left(1+\frac{1}{q}\right)(1+q)\right]^{1/5}\mathcal{M}. \tag{3.38}$$

Using Eq. (3.1), we can relate the probability of $\chi_{\rm eff}$ to the individual aligned spins χ_i:

$$P(\chi_{\rm eff} \mid q) = \iint_{-1}^{1} \mathrm{d}\chi_1 \mathrm{d}\chi_2\, \delta\left(\chi_{\rm eff} - \frac{\chi_1 + q\chi_2}{1+q}\right) P(\chi_1, \chi_2). \tag{3.39}$$

In the LIGO prior, the individual spins are uncorrelated and taken from the same distribution.

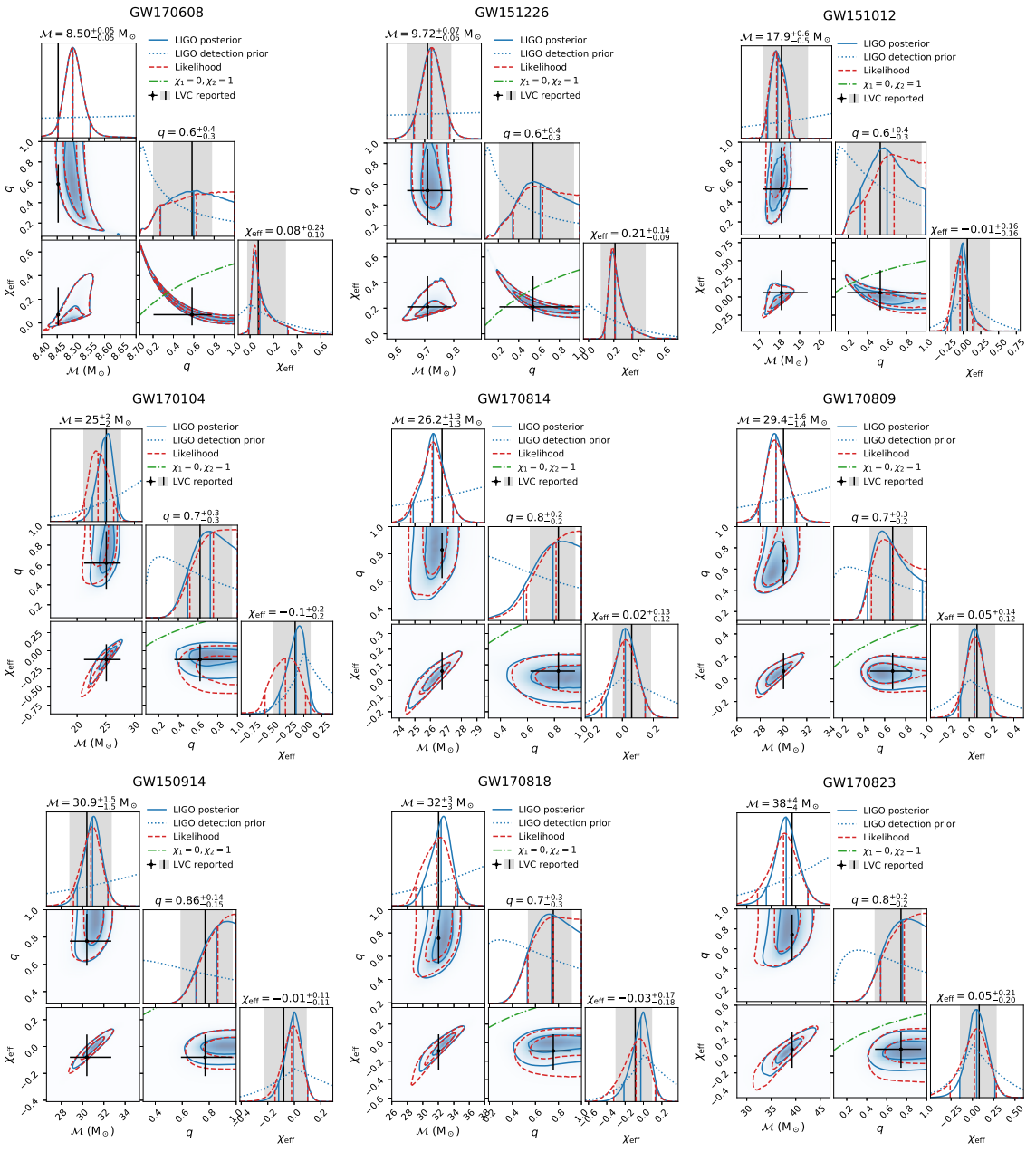


Figure 3.B.1: Parameter estimation for the BBH merger events reported to date (continued in Fig. 3.B.2). For each event, the plots on the diagonal show the marginalized likelihood (dashed line), detection prior (dotted) and posterior (solid) distributions for each parameter. The detection prior accounts for selection bias, and the likelihood is conditioned to detection. For each distribution, vertical lines show the median and minimal interval enclosing 90%. Here we use the same prior as LIGO to facilitate comparison to the reported values (black vertical line with shaded area). By "LIGO posterior" we mean the posterior distribution we computed using the LIGO prior. Off-diagonal plots show the two-dimensional marginalized likelihood and posterior. Probability contours enclosing 50% and 90% of each distribution are shown. For the q–$\chi_{\rm eff}$ plot, the case where the aligned spins of the black holes are $\chi_1 = 0, \chi_2 = 1$ is shown by a dashed-dotted line, as a proxy for what the outcome of a tidally-locked-secondary progenitor would be. The likelihood can be interpreted as the posterior distribution arising from a uniform detection prior in $\mathcal{M}, q, \chi_{\rm eff}$, so it illustrates the influence of changing the prior. The values reported by the LVC are shown by black dots with error bars. The LVC did not report the detector-frame chirp mass for the last six events (from GW170608 on), so for those cases we show their source-frame value corrected for redshift, without an uncertainty.

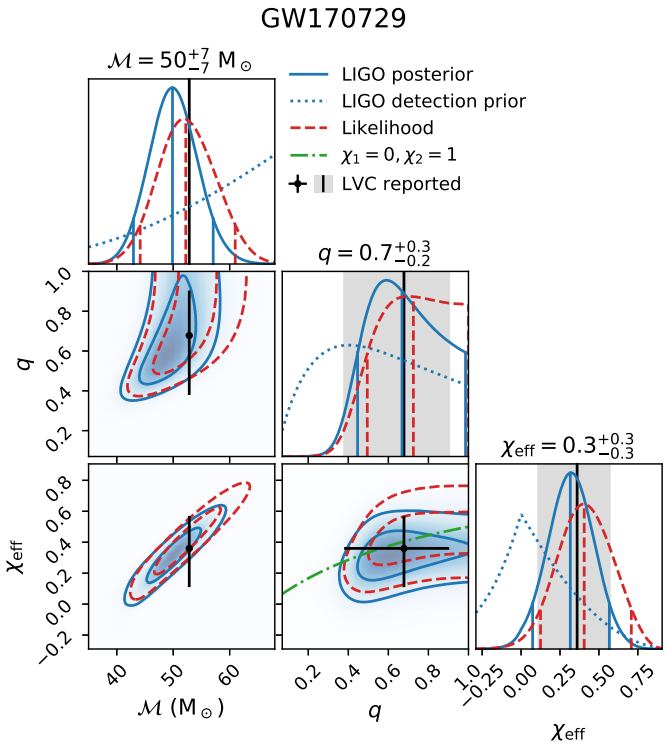


Figure 3.B.2: Continuation of Fig. 3.B.1 displaying GW170729, the heaviest and highest-χ_{eff} event.

Therefore, $P_{\mathrm{LIGO\,prior}}(\chi_1, \chi_2) = P_\chi(\chi_1) P_\chi(\chi_2)$, with

$$\begin{aligned} P_\chi(\chi) &= \int_0^1 \mathrm{d}a \int_{-1}^1 \frac{\mathrm{d}\mu}{2} \delta(\chi - a\mu) \\ &= -\frac{1}{2} \log|\chi|, \qquad |\chi| \leq 1. \end{aligned} \tag{3.40}$$

Using this, we can carry out the χ_1 integral in Eq. (3.39):

$$P_{\mathrm{LIGO\,prior}}(\chi_{\mathrm{eff}} \mid q) = (1+q) \int_a^b \mathrm{d}\chi_2 P_\chi\left((1+q)\chi_{\mathrm{eff}} - q\chi_2\right) P_\chi(\chi_2), \tag{3.41}$$

where the integration limits are

$$\begin{aligned} a &= \max\left\{\frac{(1+q)\chi_{\mathrm{eff}} - 1}{q}, -1\right\} \\ b &= \min\left\{\frac{(1+q)\chi_{\mathrm{eff}} + 1}{q}, 1\right\}. \end{aligned} \tag{3.42}$$

In practice, we compute the integral in Eq. (3.41) by quadrature (see Ng et al. (2018) for an analytical approximation). The total LIGO prior for $\mathcal{M}, q, \chi_{\mathrm{eff}}$ is given by the product of Eqs. (3.38) and (3.41).

The LIGO detection prior shown in Fig. 3.B.1 is obtained by multiplying the astrophysical prior

by the sensitive volume of the detector. We show the detection prior so that any deviations of the posterior from the prior are driven by the data and not by the parameter-dependent sensitivity of the detector.

Chapter 4

Population Inference Combining Confident and Marginal Events

This Chapter has been adapted from Roulet et al. (2020).

We perform a statistical inference of the astrophysical population of binary black hole (BBH) mergers observed during the first two observing runs of Advanced LIGO and Advanced Virgo, including events reported in the GWTC-1 and IAS catalogs. We derive a novel formalism to fully and consistently account for events of arbitrary significance. We carry out a software injection campaign to obtain a set of mock astrophysical events subject to our selection effects, and use the search background to compute the astrophysical probabilities $p_{\rm astro}$ of candidate events for several phenomenological models of the BBH population. We emphasize that values of $p_{\rm astro}$ depend on both the astrophysical and background models. Finally, we combine the information from individual events to infer the rate, spin, mass, mass-ratio and redshift distributions of the mergers. The existing population does not discriminate between random spins with a spread in the effective spin parameter, and a small but nonzero fraction of events from tidally-torqued stellar progenitors. The mass distribution is consistent with one having a cutoff at $m_{\rm max} = 41^{+10}_{-5}\,{\rm M}_\odot$, while the mass ratio favors equal masses; the mean mass ratio $\overline{q} > 0.67$. The rate shows no significant evolution with redshift. We show that the merger rate restricted to BBHs with a primary mass between 20–30 $\rm M_\odot$, and a mass ratio $q > 0.5$, and at $z \sim 0.2$, is 1.5–5.3 $\rm Gpc^{-3}yr^{-1}$ (90% c.l.); these bounds are model independent and a factor of ~ 3 tighter than that on the local rate of all BBH mergers, and hence are a robust constraint on all progenitor models. Including events from our catalog increases the Fisher information about the BBH population by $\sim 47\%$, and tightens constraints on its parameters.

4.1 Introduction

The Advanced LIGO (Aasi et al., 2015) and Advanced Virgo (Acernese et al., 2014) detectors have detected gravitational waves (GW) from a number of compact binary mergers since the start of the first observing run in 2015. Summarizing their first and second observing runs (O1 and O2), the LIGO and Virgo Collaboration (LVC) released a catalog (GWTC-1, see Ref. (Abbott et al., 2019a)) with ten BBH mergers and one binary neutron star (BNS) merger. The recently concluded third observing run (O3) has yielded a second BNS merger (Abbott et al., 2020a) and three new BBH mergers (Abbott et al., 2020b,d,c), with more to be announced.

The LVC has released the raw strain data from O1 and O2 (Abbott et al., 2021), and several independent teams have reanalyzed this dataset (Nitz et al., 2019a, 2020; Antelis and Moreno, 2019; Venumadhav et al., 2019, 2020; Zackay et al., 2019a). In previous work, we identified nine additional BBH events with significance ranging from high to marginal levels (Venumadhav et al., 2019; Zackay et al., 2019b; Venumadhav et al., 2020; Zackay et al., 2019a) (for simplicity, we will refer to these events using the abbreviation IAS, after the Institute for Advanced Study). Several of these events were independently confirmed by Ref. (Nitz et al., 2020), who also identified another significant event.

Despite being among the most detectable and accurately modeled GW sources, the origin of merging BBHs remains unclear. A variety of astrophysical formation mechanisms have been proposed, including isolated binary stellar evolution through a common envelope phase (Nelemans et al., 2001; Belczynski et al., 2002; Voss and Tauris, 2003; Belczynski et al., 2007, 2008; Dominik et al., 2013; Belczynski et al., 2014; Mennekens and Vanbeveren, 2014; Spera et al., 2015; Eldridge and Stanway, 2016; Stevenson et al., 2017b; Mapelli et al., 2017; Giacobbo et al., 2017; Mapelli and Giacobbo, 2018; Kruckow et al., 2018; Giacobbo and Mapelli, 2018), chemically homogeneous stellar evolution (Marchant et al., 2016; de Mink and Mandel, 2016; Mandel and de Mink, 2016), or dynamic capture and hardening of binaries in dense stellar systems such as globular clusters (Zwart and McMillan, 1999; O'Leary et al., 2006; Sadowski et al., 2008; Downing et al., 2010, 2011; Samsing et al., 2014; Rodriguez et al., 2015, 2016a; Askar et al., 2016), nuclear clusters (Antonini and Rasio, 2016; Petrovich and Antonini, 2017), and young open clusters (Ziosi et al., 2014; Mapelli, 2016; Banerjee, 2017; Chatterjee et al., 2017). Alternatively, mergers can be prompted by interactions with gas and stars in AGN disks (McKernan et al., 2012; Stone et al., 2016; Bartos et al., 2017), or through the Kozai–Lidov effect in the presence of a supermassive black hole (Antonini and Perets, 2012) or in triple (or higher multiplicity) systems (Antonini et al., 2014; Kimpson et al., 2016; Antonini et al., 2017; Liu

and Lai, 2018; Hamers et al., 2015).

The growing number of detections has spawned many efforts to statistically characterize the population of these systems, with the main goal of comparing the observed sample statistics with the predictions of different proposed formation channels. The distributions of the BBH merger rate, masses, spins and redshifts have been studied in the literature (Vitale et al., 2017b; Talbot and Thrane, 2017; Abbott et al., 2016c; Fishbach and Holz, 2017; Hotokezaka and Piran, 2017; Fishbach et al., 2018; Farr et al., 2018; Wysocki et al., 2018, 2019; Roulet and Zaldarriaga, 2019; Abbott et al., 2019b; Fishbach and Holz, 2020). Several of these works were based on the entirety or a subset of the ten confident BBH detections reported in GWTC-1. Other recent works also included events from the IAS catalog in the BBH population analyses (Piran and Piran, 2020; Gayathri et al., 2020; Galaudage et al., 2020; Pratten and Vecchio, 2020). However, when considering the events as a population, we have to appropriately account for the marginal significance of some of these events: for such events, the probability of astrophysical origin itself can depend on the population model being considered (this was previously noted in Ref. (Nitz et al., 2020)). Moreover, the probability of astrophysical origin for a candidate depends on the search pipeline that found it, since the definition involves the levels of comparable foreground and background triggers, subject to the same selection effects. Hence, the sensitivity of the search pipeline is a crucial piece of information when inferring astrophysical event rates and correcting for selection effects. In this work, we inject synthetic signals into the O1 and O2 runs to empirically measure the spacetime-volume our pipeline is sensitive to.

In this paper, we develop a general framework to treat the problem of population inference using detections of arbitrary significance, and apply it in an analysis that accounts for the BBH events in the GWTC-1 and IAS catalogs. This problem was first studied by Gaebel et al. (2019), who introduced a formalism for population inference using a mixture of noise and signal triggers. They demonstrated its implementation for a simulation on a simplified parameter space with analytical prescriptions for the foreground and background distributions. Our work expands on this by developing a framework that can be cast in terms of a small number of quantities that are straightforward to compute, and thus more amenable to real-data applications, and implementing it on the O1 and O2 datasets. Galaudage et al. (2020) developed a different implementation and applied it to include the events in the IAS catalog. We address some issues with this treatment, and how our formalism deals with them, in Appendix 4.C.

Three BBH mergers from the O3 run have been recently reported (Abbott et al., 2020b,d,c). We exclude them from the present analysis since as highlights from a (yet unreleased) O3 catalog, they cannot be simply added to a set of homogeneously selected events for population inference.

We organize the rest of the paper as follows: in §4.2, we derive the likelihood of a population model as a function of events of arbitrary significance, and then we present the algorithm to compute in practice the various quantities involved. In addition, we estimate the gain in the Fisher information from the inclusion of marginal triggers, and use it as a guide to set a convenient threshold for which triggers to include in the analysis. In §4.3, we report new constraints on the astrophysical population of BBH mergers. In particular, we study the merger rate distribution's dependence on spin, mass, mass ratio and source redshift. Finally, we draw conclusions in §4.4. We quantify the sensitive volume–time of our search pipeline by means of software injection in Appendix 4.A. We describe technical details of computing the astrophysical probability, $p_{\rm astro}$, in Appendix 4.B. We address the relation to previous related work in Appendix 4.C. We provide evidence that our method is robust to importance sampling stochastic errors in Appendix 4.D.

4.2 Framework

In this section we derive the likelihood function for the set of triggers above a given threshold in a pipeline due to a distribution of BBHs, present an algorithm for its practical evaluation, and derive the amount of Fisher information contained in the marginal triggers. The likelihood constrains which population models are consistent with the data. We will follow the notation of Mandel et al. (2019), see also Refs. (Thrane and Talbot, 2019; Vitale, 2020) for an introductory treatment.

4.2.1 Model likelihood

As a preliminary matter, we define the data as the observable quantities that detectors output, along with any quantities derived from this that we use in the search. A datum d consists of (a) a measured strain timeseries at each detector, long enough to capture all astrophysical information available in a putative signal, and (b) derived quantities, e.g. detection statistics or statistics used for signal quality tests. Note that this excludes BBH parameters such as masses and spins, which are not directly observable. The full dataset (here, O1 and O2) can be thought of a large set of points in this high-dimensional space, one for every datum. These realizations contain detector noise plus, in comparatively very few cases, astrophysical signals.

We restrict the analysis to a small set of selected data realizations $\{d_i\}$ ("triggers") defined under some criteria such that each excluded individual datum is very unlikely to contain an astrophysical signal. This set may contain both secure and marginal events. We will select the triggers using the search pipeline described in Ref. (Venumadhav et al., 2019) and a threshold on its detection statistic

(henceforth "detection threshold", we discuss our choice in §4.2.4); it is the task of the search pipeline to compute the detection statistic for all data. We will assume triggers to be independent: since triggers are rare, the fact that a trigger uses a sample does not affect significantly the total number of available samples for other triggers to happen. The assumption of independence could be invalid, though, if different triggers of astrophysical origin were produced by multiple images of a single gravitationally lensed source (Dai et al., 2020); we will not consider such possibility in this analysis.

Under this assumption, the search for BBHs in the full dataset is a Poisson process that generates a set $\{d_i\}$ of $N_{\rm trig}$ triggers in the above-threshold subregion of the phase space of data. The likelihood of observing this set of triggers is given by the Poisson distribution

$$P(N_{\rm trig}, \{d_i\} \mid \lambda) = \frac{e^{-N_a(\lambda)-N_b}}{N_{\rm trig}!} \prod_{i=1}^{N_{\rm trig}} \left(\left.\frac{\mathrm{d}N_a}{\mathrm{d}d}\right|_{d_i}(\lambda) + \left.\frac{\mathrm{d}N_b}{\mathrm{d}d}\right|_{d_i} \right), \tag{4.1}$$

where $N_a(\lambda)$ is the expected number of triggers in the set with astrophysical origin under a model for the source population described by a set of parameters λ, N_b is the expected number of noise background triggers in the set, and the terms $\mathrm{d}N_a/\mathrm{d}d, \mathrm{d}N_b/\mathrm{d}d$ are the rate densities for triggers under the astrophysical and background hypotheses.

We express the expected rate density of astrophysical triggers in terms of the physical merger rate through

$$\left.\frac{\mathrm{d}N_a}{\mathrm{d}d}\right|_{d}(\lambda) = \int \mathrm{d}\theta\, P(d \mid \theta)\, \frac{\mathrm{d}N_a}{\mathrm{d}\theta}(\theta \mid \lambda), \tag{4.2}$$

where θ are a set of parameters that characterize each merger (e.g. masses, spins, distance, sky position, orbital orientation, time, etc.) and $P(d \mid \theta)$ is the parameter likelihood. For triggers that pass all signal quality tests, a Gaussian noise model is typically a good description of the parameter likelihood:

$$P(d \mid \theta) \propto \exp\left(-\frac{1}{2}\langle d - h(\theta) \mid d - h(\theta)\rangle\right), \tag{4.3}$$

where $h(\theta)$ is the GW strain model and, in a slight abuse of notation, d is the measured strain. As is standard in the GW literature (Thorne, 1987), the argument of the exponential in Eq. (4.3) is the inverse-variance weighted inner product between two real-valued time series x and y,

$$\langle x \mid y\rangle = 4\,\mathrm{Re}\int_0^\infty \mathrm{d}f\, \frac{\tilde{x}^*(f)\,\tilde{y}(f)}{S_n(f)}, \tag{4.4}$$

where $S_n(f)$ is the one-sided power spectral density (PSD) of the detector noise, tildes indicate Fourier transforms, and asterisks complex conjugation.

The expected rate density of background triggers also depends on the detection pipeline. We will estimate it using the method of timeslides, see Appendix 4.B for further details.

The expected number of astrophysical triggers is

$$\begin{aligned} N_a(\lambda) &= \int_{d>\mathrm{th}} \mathrm{d}d \left.\frac{\mathrm{d}N_a}{\mathrm{d}d}\right|_d (\lambda) \\ &= \int \mathrm{d}\theta \, \frac{\mathrm{d}N_a}{\mathrm{d}\theta}(\theta \mid \lambda)\, p_{\mathrm{det}}(\theta). \end{aligned} \tag{4.5}$$

In the first line, the d integral runs over all data realizations that would result in a pipeline statistic above the detection threshold. In the second line, we introduced the detection efficiency for a source with parameters θ:

$$p_{\mathrm{det}}(\theta) = \int_{d>\mathrm{th}} \mathrm{d}d \, P(d \mid \theta), \tag{4.6}$$

The probability $p_{\mathrm{astro},i}$ that the ith trigger is of astrophysical origin depends on models for both astrophysical events and noise triggers. In this work, we will fix the model for noise triggers, but vary the astrophysical one. Given an astrophysical model described by parameters λ,

$$p_{\mathrm{astro},i}(\lambda) = \left.\frac{\mathrm{d}N_a(\lambda)}{\mathrm{d}N_a(\lambda) + \mathrm{d}N_b}\right|_{d_i}. \tag{4.7}$$

For practical evaluation, we rescale Eq. (4.1) into the following form while keeping the dependence on the model parameters λ,

$$\begin{aligned} P(N_{\mathrm{trig}}, \{d_i\} \mid \lambda) &\propto \frac{P(N_{\mathrm{trig}}, \{d_i\} \mid \lambda)}{P(N_{\mathrm{trig}}, \{d_i\} \mid \lambda_0)} \\ &\propto e^{-N_a(\lambda)} \prod_{i=1}^{N_{\mathrm{trig}}} \left.\frac{\mathrm{d}N_a(\lambda) + \mathrm{d}N_b}{\mathrm{d}N_a(\lambda_0) + \mathrm{d}N_b}\right|_{d_i} \\ &= e^{-N_a(\lambda)} \prod_{i=1}^{N_{\mathrm{trig}}} \left[\left.\frac{\mathrm{d}N_a(\lambda)}{\mathrm{d}N_a(\lambda_0)}\right|_{d_i} \left.\frac{\mathrm{d}N_a(\lambda_0)}{\mathrm{d}N_a(\lambda_0) + \mathrm{d}N_b}\right|_{d_i} + \left.\frac{\mathrm{d}N_b}{\mathrm{d}N_a(\lambda_0) + \mathrm{d}N_b}\right|_{d_i} \right] \\ &= e^{-N_a(\lambda)} \prod_{i=1}^{N_{\mathrm{trig}}} \left[\left.\frac{\mathrm{d}N_a(\lambda)}{\mathrm{d}N_a(\lambda_0)}\right|_{d_i} p_{\mathrm{astro},i}(\lambda_0) + \big(1 - p_{\mathrm{astro},i}(\lambda_0)\big) \right], \end{aligned} \tag{4.8}$$

where λ_0 corresponds to a fiducial source population model that we are free to choose. Equation (4.8) converges to a meaningful number as one relaxes the detection threshold and includes arbitrarily insignificant triggers with $p_{\mathrm{astro}} \to 0$ (we will discuss this point further in §4.2.4). In the opposite limit in which we only include events with absolute certainty of astrophysical origin ($p_{\mathrm{astro}} = 1$), we

recover the standard result (e.g. in the notation of Mandel et al. (2019))

$$\begin{aligned} P(N_{\rm trig}, \{d_i\} \mid \lambda, p_{{\rm astro},i}=1) &= \frac{e^{-N_a(\lambda)}}{N_{\rm trig}!} \prod_{i=1}^{N_{\rm trig}} \left.\frac{{\rm d}N_a}{{\rm d}d}\right|_{d_i}(\lambda) \\ &= \frac{e^{-N_a(\lambda)}}{N_{\rm trig}!} \left[N_a(\lambda)\right]^{N_{\rm trig}} \prod_{i=1}^{N_{\rm trig}} \frac{\int {\rm d}\theta\, P(d_i \mid \theta) \frac{{\rm d}N_a}{{\rm d}\theta}(\theta \mid \lambda)}{N_a(\lambda)} \\ &= P(N_{\rm trig} \mid \lambda) \prod_{i=1}^{N_{\rm trig}} \frac{\int {\rm d}\theta\, P_{\rm pop}(\theta \mid \lambda')\, P(d_i \mid \theta)}{\int {\rm d}\theta\, P_{\rm pop}(\theta \mid \lambda')\, p_{\rm det}(\theta)}, \end{aligned} \tag{4.9}$$

where $P(N_{\rm trig} \mid \lambda)$ follows standard Poisson statistics (this can also be obtained by marginalizing over $\{d_i\}$ in Eq. (4.1)). In the above formula, λ' are the population parameters that characterize the shape of the un-normalized astrophysical distribution $P_{\rm pop}$, separated out from an overall merger rate, which we will call R. The overall rate R cancels inside the product since both ${\rm d}N_a/{\rm d}\theta$ and N_a are linearly proportional to it. However, note that once we include events with $0 < p_{\rm astro} < 1$, the value of $p_{\rm astro}(\lambda)$ depends on the rate even at fixed population shape and such a clean separation does not occur.

Nevertheless, we can exploit the linear dependence of $N_a(\lambda)$ and ${\rm d}N_a/{\rm d}\theta$ on the rate parameter to evaluate these terms efficiently. We make explicit the decomposition of the population parameters λ into rate R and shape λ':

$$\frac{{\rm d}N_a}{{\rm d}\theta}(\theta \mid \lambda) = R\, f(\theta \mid \lambda'), \tag{4.10}$$

where $f(\theta \mid \lambda')$ is normalized according to

$$\lim_{VT \to 0} \frac{1}{VT} \int_{VT} {\rm d}\theta\, f(\theta \mid \lambda') = 1 \tag{4.11}$$

in the local Universe, so that R is the local merger rate per unit time per unit volume. Note that the source distance and the arrival time of the signal are among the parameters θ, and for these we do not normalize their distribution to integrate to unity over some domain since they do not have a natural scale. From Eqs. (4.2) and (4.10), the astrophysical number density ratio for the ith trigger in Eq. (4.8) is

$$\left.\frac{{\rm d}N_a(\lambda)}{{\rm d}N_a(\lambda_0)}\right|_{d_i} = \frac{R}{R_0}\, w_i(\lambda'; \lambda'_0), \tag{4.12}$$

where we define

$$w_i(\lambda'; \lambda'_0) := \frac{\int {\rm d}\theta\, P(d_i \mid \theta)\, f(\theta \mid \lambda')}{\int {\rm d}\theta\, P(d_i \mid \theta)\, f(\theta \mid \lambda'_0)}. \tag{4.13}$$

Likewise, the expected detection rate is

$$N_a(\lambda) = R \cdot \overline{VT}(\lambda'), \tag{4.14}$$

where, in accord with Eq. (4.5),

$$\overline{VT}(\lambda') = \int \mathrm{d}\theta \, f(\theta \mid \lambda') \, p_{\mathrm{det}}(\theta) \tag{4.15}$$

is the population-averaged sensitive volume–time of the detector network (analogous to $\alpha(\lambda')$ of Mandel et al. (2019) but with a different normalization choice; equivalent to $\mathcal{V}(\Lambda)$ of Galaudage et al. (2020)). $\overline{VT}(\lambda')$ depends on the search pipeline and detection threshold used.

In terms of these quantities, Eq. (4.8) takes the form

$$P(N_{\mathrm{trig}}, \{d_i\} \mid \lambda) \propto e^{-R\,\overline{VT}(\lambda')} \prod_{i=1}^{N_{\mathrm{trig}}} \left[\frac{R}{R_0} \, w_i(\lambda'; \lambda'_0) \, p_{\mathrm{astro},i}(\lambda_0) + (1 - p_{\mathrm{astro},i}(\lambda_0)) \right]. \tag{4.16}$$

4.2.2 Likelihood evaluation

In order to evaluate Eq. (4.16), we need to evaluate three types of terms: $w_i(\lambda'; \lambda'_0)$, $\overline{VT}(\lambda')$ and $p_{\mathrm{astro},i}(\lambda_0)$.

We estimate $w_i(\lambda'; \lambda'_0)$ from the integral in Eq. (4.13) using a Monte Carlo method:

$$w_i(\lambda'; \lambda'_0) \approx \mathcal{W}_i(\lambda'; \lambda'_0) := \frac{\displaystyle\sum_{j=1}^{S_i} \frac{1}{\pi(\theta^i_j)} \, f(\theta^i_j \mid \lambda')}{\displaystyle\sum_{j=1}^{S_i} \frac{1}{\pi(\theta^i_j)} \, f(\theta^i_j \mid \lambda'_0)}, \tag{4.17}$$

where $\{\theta^i_j : j = 1, \ldots, S_i\}$ are samples from the posterior distribution of the parameters for the ith trigger, obtained under a prior $\pi(\theta)$.

Similarly, we can evaluate $\overline{VT}(\lambda')$ by reweighting a set of injections (Tiwari, 2018). We add synthetic signals to the data in software and run the detection pipeline (with vetoes and choices that are as close as possible to those in the 'production' run[1]) to determine which injections would have been found. From Eq. (4.15) we can construct the estimator

$$\overline{VT}(\lambda') \approx \overline{\mathcal{VT}}(\lambda') := \frac{1}{N_{\mathrm{inj}}} \sum_{j > \mathrm{th}} \frac{f(\theta_j \mid \lambda')}{P(\theta_j \mid \lambda'_{\mathrm{inj}})}. \tag{4.18}$$

[1]We change some choices in order to keep computational cost manageable, as detailed in Appendix 4.A.1.

Here, we denote by $P(\theta \mid \lambda'_{\rm inj})$ the distribution of source parameters from which we generate injections. Note that $N_{\rm inj}$ is the total (found and missed) number of injections, but the sum runs only over those above the detection threshold.

Finally, we have to compute the reference $p_{\rm astro}(\lambda_0)$ for all triggers under consideration. Note that these correspond to a particular astrophysical model λ_0, so we cannot use the numbers reported by a pipeline at face value without regard to λ_0. In §4.2.3 we describe our choice of λ_0. According to Eq. (4.7), $p_{\rm astro}$ requires estimating both the foreground and the background of the search pipeline. We estimate the foreground using injections and the background using timeslides (see Appendix 4.B for details). We report the values of $p_{\rm astro}(\lambda_0)$ in Table 4.3.1.

As long as the population models λ of interest do not have too many parameters, we can evaluate the rate-independent estimators $\{\mathcal{W}_i(\lambda';\lambda'_0)\}$ and $\overline{\mathcal{VT}}(\lambda')$ on an auxiliary λ' grid. We then use them to evaluate Eq. (4.16) on a λ grid, that incorporates the dependence with rate avoiding redundant reevaluation of $\{\mathcal{W}_i(\lambda';\lambda'_0)\}$ and $\overline{\mathcal{VT}}(\lambda')$. Note that this procedure extends readily to a situation where $\mathrm{d}N_a$ depends linearly on multiple population parameters, a commonly encountered case being the branching ratios of a "mixture model" which consists of a linear combination of several subpopulations.

The BBH merger GW170608 occurred while one LIGO detector was not in nominal observing mode (Abbott et al., 2017c). Data from such periods are not publicly accessible so our injections do not simulate this type of events. Although we cannot use GW170608 to inform the astrophysical rate, the event contains valuable information about the shape parameters λ'. In order to include GW170608 consistently in our analysis, we single it out as an additional event with $p_{\rm astro} = 1$ that is not counted in $N_a(\lambda)$. We replace Eq. (4.8) by

$$P(\{d_i\}, N_{\rm trig} \mid \lambda) = P(\{d_{i\neq \rm GW170608}\}, N_{\rm trig} - 1 \mid \lambda)\, P(d_{\rm GW170608} \mid \lambda'), \tag{4.19}$$

where we choose to normalize

$$P(d_{\rm GW170608} \mid \lambda') = \frac{1}{N_a(p_{\rm astro} \approx 1)} \left.\frac{\mathrm{d}N_a}{\mathrm{d}d}\right|_{\rm GW170608} \tag{4.20}$$

so that it integrates to 1 over all non-observing-mode data realizations that would have yielded a $p_{\rm astro} \approx 1$ event. We implement Eq. (4.19) under the approximation $N_a(p_{\rm astro} \approx 1) \propto N_a(\lambda)$, which in practice amounts to dividing Eq. (4.8) by $N_a(\lambda)$ (and in general, by $[N_a(\lambda)]^n$ to include n additional events).

4.2.3 Choice of reference and injection distributions

The Monte Carlo estimators $\{\mathcal{W}_i(\lambda';\lambda'_0)\}$ and $\overline{\mathcal{VT}}(\lambda')$ (Eqs. (4.17) and (4.18)) are unbiased, in the sense that their expectation values are

$$\langle \mathcal{W}_i(\lambda';\lambda'_0)\rangle = w_i(\lambda';\lambda'_0) \tag{4.21}$$

$$\langle \overline{\mathcal{VT}}(\lambda')\rangle = \overline{VT}(\lambda') \tag{4.22}$$

regardless of the choice of reference and injection populations, λ'_0 and $\lambda'_{\rm inj}$. However, due to the finite number of samples used, they have variances that depend on the choice of λ'_0 and $\lambda'_{\rm inj}$. In this subsection, we discuss choices that allow robust estimation of $w_i(\lambda';\lambda'_0)$ and $\overline{VT}(\lambda')$.

The main requirement for these importance sampling estimators is that the proposal distribution from which the samples are taken does not vanish at places where the integrand (target distribution) is nonzero, lest the reweighting become pathological. Thus, neither the parameter estimation prior $\pi(\theta)$ nor the reference population $f(\theta \mid \lambda'_0)$ in Eq. (4.17) should vanish anywhere the likelihood $P(d_i \mid \theta)$ has support. In general, the estimator variance will be smaller when the proposal distribution more closely matches the target distribution.

The effective spin

$$\chi_{\rm eff} := \frac{\chi_{1z} + q\chi_{2z}}{1+q} \tag{4.23}$$

is the spin variable that can be measured best, where $q = m_2/m_1$ is the mass ratio and χ_{1z}, χ_{2z} are the dimensionless spin components in the direction of the binary's orbital angular momentum. Spin components in the orbital plane are harder to measure than and no evidence for or against them has been found for any of the mergers of O1 and O2. The same is true for gravitational radiation modes beyond the quadrupolar $(\ell, |m|) = (2, 2)$. To simplify the analysis, we will neglect in-plane spin components and higher-order modes in the following, and we do not expect that this will change the results significantly. In particular, we will use the aligned-spin, quadrupolar radiation waveform approximant `IMRPhenomD` (Khan et al., 2016) to model gravitational wave signals throughout this work. For parameter estimation, we use the relative binning method for likelihood evaluation (Zackay et al., 2018) coupled to the `PyMultinest` sampler (Buchner et al., 2014).

We adopt a parameter estimation prior $\pi(\theta)$ that is uniform in detector-frame masses, effective spin $\chi_{\rm eff}$ and luminosity volume. Some events, most notably GW151216 (Zackay et al., 2019b), have high effective spins, so a prior that does not vanish for extreme values of $\chi_{\rm eff}$ is convenient. We parametrize the two spins in terms of the well measured $\chi_{\rm eff}$ and a poorly measured variable

$\chi_{\rm diff} \coloneqq (q\chi_{1z} - \chi_{2z})/(1+q)$ that controls how much either binary component contributes to $\chi_{\rm eff}$. We implement the spin prior as flat in $\chi_{\rm eff}$, and flat in $\chi_{\rm diff}$ conditioned on $\chi_{\rm eff}$ within the Kerr bound $|\chi_{1,2}| < 1$.

For the reference population $f(\theta \mid \lambda_0')$, we narrow our focus to $\theta = \{m_{1\rm s}, q, \chi_{\rm eff}, D_L\}$, which are the measurable variables that lack a natural prior, and adopt the following (factorized) joint distribution: truncated power-law in the primary source-frame mass $m_{1\rm s}$ and uniformity in q, $\chi_{\rm eff}$ and luminosity volume, with λ_0' that lie in the bulk of reported posterior distributions (see e.g. Abbott et al., 2019b):

$$f(m_{1\rm s}, q, \chi_{\rm eff}, D_L \mid \lambda_0') \propto m_{1\rm s}^{-\alpha_0} D_L^2 \tag{4.24}$$

for $m_{1\rm min} < m_{1\rm s} < m_{1\rm max}$ and $q_{\rm min} < q < 1$, with $\alpha_0 = 2.35$, $m_{1\rm min} = 3\,{\rm M}_\odot$, $m_{1\rm max} = 120\,{\rm M}_\odot$ and $q_{\rm min} = 1/20$. These ranges are broad enough to encompass the likelihood support of all the triggers we include. Throughout this work we will implicitly assume a uniform distribution for $\chi_{\rm diff}$, arrival time, orbital phase, orbital orientation and sky position. When computing $p_{\rm astro}(\lambda_0)$ we will use a fiducial rate $R_0 = 10^{1.5}\,{\rm Gpc}^{-3}\,{\rm yr}^{-1}$. Note that the choice of λ_0 does not affect the final results (provided the reweighing process presents no pathologies, as we demonstrate in Appendix 4.D) and therefore using previous analyses to inform our choice does not bias our conclusions.

Finally, we choose an injection distribution (Eq. (4.18)) that approximately matches the integrand in Eq. (4.15). We adopt

$$P(\theta \mid \lambda_{\rm inj}') = Z^{-1}\, f(\theta \mid \lambda_0')\, \hat{p}_{\rm det}(\theta), \tag{4.25}$$

$$\begin{aligned} Z &= \int {\rm d}\theta\, f(\theta \mid \lambda_0')\, \hat{p}_{\rm det}(\theta) \\ &\approx 7.1\,{\rm Gpc}^3\, T_{\rm obs}, \end{aligned} \tag{4.26}$$

where $\hat{p}_{\rm det}(\theta)$ is some semianalytical approximation of $p_{\rm det}(\theta)$, $T_{\rm obs}$ is the observation time span on which the injections are made, and the value of Z reported in Eq. (4.26) corresponds to the choices that follow. We use

$$\hat{p}_{\rm det}(\theta) = \int_{\rho^2_{\rm th}}^{\infty} {\rm d}\rho^2 \chi^2(\rho^2, 10, \rho^2_*(\theta)). \tag{4.27}$$

Here, we use $\rho^2_{\rm th} = 60$ as an approximate detection threshold. We define the expected signal-to-noise ratio (SNR) of an event with parameters θ:

$$\rho_*(\theta) = \rho_{1\rm Mpc}(m_1^{\rm det}, q, \chi_{1z}, \chi_{2z})\, A(\alpha, \delta, \iota, \psi, t)\, \frac{1\,{\rm Mpc}}{D_L}. \tag{4.28}$$

This is computed for a fiducial noise PSD, which we define in each frequency bin as the 10th percentile of 200 random 4096 s segments of Hanford and Livingston O2 data. In addition, $A = \sqrt{A_{\mathrm{H}}^2 + A_{\mathrm{L}}^2}$ is the Hanford–Livingston antenna pattern, where the angular factors $0 < A_{\mathrm{H,L}} < 1$ can be found, e.g., in Ref. (Sathyaprakash and Schutz, 2009). We define $\rho_{1\mathrm{Mpc}} = \langle h \mid h \rangle^{1/2}$ to be the single-detector SNR of an optimally oriented source at a fiducial distance of 1 Mpc, which we interpolate on a grid of intrinsic parameters.

The non-central chi-squared distribution in Eq. (4.27) models the distribution of SNR^2 recovered by a search pipeline for a signal with parameters θ in the presence of Gaussian noise and maximized over ten degrees of freedom (Jaranowski and Królak, 2012). Six degrees of freedom model maximization by the pipeline over the signal amplitude, phase and arrival time independently in the two LIGO detectors. The remaining four model maximization over template parameters (our waveform templates are elements of a metric space of up to 4 dimensions (Roulet et al., 2019)). Eq. (4.27) neglects a variety of effects present in the search process, such as template bank discreteness and boundaries, signal coherence across detectors, detector sensitivity variations, noise non-stationarity and non-Gaussianity, and signal quality vetoes. This is acceptable, since $\hat{p}_{\mathrm{det}}$ is only used for choosing a convenient injection distribution: all these effects are accounted for by the injections as per Eq. (4.18). Moreover, they make $\hat{p}_{\mathrm{det}}$ a somewhat optimistic estimate of p_{det}, which is desirable as it makes the proposal distribution (Eq. (4.25)) broader than the target distribution (Eq. (4.15)).

We generate the set of source parameters $\{\theta_j\}$ for injected signals by drawing samples from the distribution in Eq. (4.25) with the `PyMultinest` sampler (Buchner et al., 2014), with which we simultaneously evaluate the normalization constant Z as reported in Eq. (4.26). Eqs. (4.18) and (4.25) yield

$$\overline{\mathcal{VT}}(\lambda') = \frac{Z}{N_{\mathrm{inj}}} \sum_{j>\mathrm{th}} \frac{f(\theta_j \mid \lambda')}{f(\theta_j \mid \lambda'_0)\, \hat{p}_{\mathrm{det}}(\theta_j)}. \tag{4.29}$$

We report technical details and results from the injection campaigns in Appendix 4.A.

4.2.4 Choice of detection threshold

We now derive a criterion for the choice of the threshold at which a trigger is deemed sufficiently informative to be included in the analysis. To that end we compute the Fisher information as a function of detection threshold. This also serves as an estimate of the amount of information about the BBH population gained by including marginal events in the present analysis.

The information that the data carry about the BBH distribution is encoded in the Fisher matrix

$$I(\lambda)_{mn} = -\left\langle \partial^2_{mn} \log P(\{d_i\}, N_{\rm trig} \mid \lambda) \right\rangle_{\{d_i\}, N_{\rm trig}}. \tag{4.30}$$

Here, $\partial^2_{mn} = \partial^2/\partial\lambda_m \partial\lambda_n$ is the second derivative with respect to population parameters λ_m and λ_n, and the subscript denotes that the expectation value is over the distribution of observations (Eq. (4.1)):

$$P(\{d_i\}, N_{\rm trig} \mid \lambda) = P(N_{\rm trig} \mid \lambda) \prod_{i=1}^{N_{\rm trig}} P(d_i \mid \lambda), \tag{4.31}$$

with

$$P(N_{\rm trig} \mid \lambda) = \frac{e^{-N(\lambda)}[N(\lambda)]^{N_{\rm trig}}}{N_{\rm trig}!}, \tag{4.32}$$

$$P(d_i \mid \lambda) = \frac{1}{N(\lambda)} \frac{{\rm d}N(\lambda)}{{\rm d}d}\bigg|_{d_i}, \tag{4.33}$$

$$N(\lambda) = N_a(\lambda) + N_b. \tag{4.34}$$

N is the expected number of astrophysical and background triggers above detection threshold, and hence depends on the threshold choice. Below we quantify how this choice affects $I(\lambda)_{mn}$. Eqs. (4.30), (4.1) and (4.34) yield

$$\begin{aligned} I(\lambda)_{mn} &= -\left\langle -\partial^2_{mn} N + \sum_{i=1}^{N_{\rm trig}} \partial^2_{mn} \log \frac{{\rm d}N}{{\rm d}d}\bigg|_{d_i} \right\rangle_{\{d_i\}, N_{\rm trig}} \\ &= \partial^2_{mn} N - \left\langle \sum_{i=1}^{N_{\rm trig}} \frac{\partial^2_{mn} {\rm d}N/{\rm d}d}{{\rm d}N/{\rm d}d} \right\rangle + \left\langle \sum_{i=1}^{N_{\rm trig}} \left(\partial_m \log \frac{{\rm d}N}{{\rm d}d}\right)\left(\partial_n \log \frac{{\rm d}N}{{\rm d}d}\right) \right\rangle. \end{aligned} \tag{4.35}$$

Using Eqs. (4.32) and (4.33) we evaluate the first sum:

$$\begin{aligned} \left\langle \sum_{i=1}^{N_{\rm trig}} \frac{\partial^2_{mn} {\rm d}N/{\rm d}d}{{\rm d}N/{\rm d}d} \right\rangle &= \langle N_{\rm trig} \rangle \int {\rm d}d \frac{1}{N} \frac{{\rm d}N}{{\rm d}d} \frac{\partial^2_{mn} {\rm d}N/{\rm d}d}{{\rm d}N/{\rm d}d} \\ &= \partial^2_{mn} \int {\rm d}d \frac{{\rm d}N}{{\rm d}d} \\ &= \partial^2_{mn} N. \end{aligned} \tag{4.36}$$

Further, from Eqs. (4.34) and (4.7) we obtain

$$\begin{aligned}\partial_m \log \frac{\mathrm{d}N}{\mathrm{d}d} &= \frac{\mathrm{d}N_a}{\mathrm{d}N}\,\partial_m \log \frac{\mathrm{d}N_a}{\mathrm{d}d} \\ &= p_{\mathrm{astro}}(d)\,\partial_m \log \frac{\mathrm{d}N_a}{\mathrm{d}d},\end{aligned} \tag{4.37}$$

which we use to evaluate the second sum in Eq. (4.35). Equations (4.35), (4.36) and (4.37) yield

$$I(\lambda)_{mn} = \left\langle \sum_{i=1}^{N_{\mathrm{trig}}} p^2_{\mathrm{astro},i} \left(\partial_m \log \frac{\mathrm{d}N_a}{\mathrm{d}d}\right)\left(\partial_n \log \frac{\mathrm{d}N_a}{\mathrm{d}d}\right)\right\rangle. \tag{4.38}$$

The special case where $\lambda_m = \lambda_n = R$ is the astrophysical merger rate is particularly simple, because $\mathrm{d}N_a/\mathrm{d}d$ is proportional to the rate:

$$\frac{\partial}{\partial R} \log \frac{\mathrm{d}N_a}{\mathrm{d}d} = \frac{1}{R}, \tag{4.39}$$

so the Fisher information about the rate is

$$I(R) = \frac{1}{R^2}\left\langle \sum_{i=1}^{N_{\mathrm{trig}}} p^2_{\mathrm{astro},i} \right\rangle. \tag{4.40}$$

In general, the information each trigger carries about the population depends on its parameters and the population model through the terms in parentheses in Eq. (4.38), weighted by its p^2_{astro}. The threshold should therefore be set such that the summed p^2_{astro} of excluded triggers is much smaller than that of included triggers, while keeping their number manageable.

As an illustrative example, in Fig. 4.2.1 we study a simple toy model in which the only parameter measured on triggers is the SNR ρ, and there is a population of signals with a power-law distribution $\mathrm{d}N_a/\mathrm{d}\rho^2 \propto \rho^{-5}$ in a Gaussian background $\mathrm{d}N_b/\mathrm{d}\rho^2 \propto \exp\left(-\rho^2/2\right)$, intended to qualitatively describe features of GW signals (Schutz, 2011) and detector noise. We consider the problem of inferring the astrophysical rate R, that is, the normalization of the power-law component in this model. For Fig. 4.2.1 we adopt fiducial normalizations so that there are respectively $N_a(\rho^2 > 65) = 15$ and $N_b(\rho^2 > 65) = 1$ expected foreground and background events with a ρ^2 louder than 65, roughly matching the numbers observed in O1 and O2. The ratio between the foreground and background distributions determines $p_{\mathrm{astro}}(\rho^2)$ through Eq. (4.7), which in turn allows us to compute the Fisher information $I(R)$ with Eq. (4.40). In the top panel of Fig. 4.2.1 we show the Fisher information as a function of a detection threshold ρ^2_{th} above which triggers are included in the likelihood, Eq. (4.8). We find that the information contained in the faint triggers is limited: even though there are many faint signals, their contribution to $I(R)$ is strongly suppressed by p^2_{astro}, shown in the bottom panel.

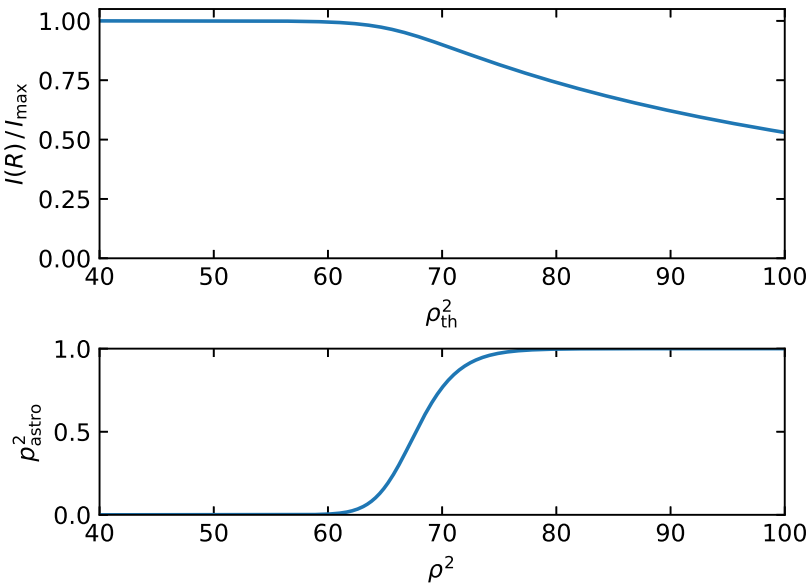


Figure 4.2.1: *Top panel:* Fisher information about rate as a function of detection threshold, in a toy model where there is a foreground power-law in SNR with unknown rate and a Gaussian background. The information saturates as the threshold is lowered, thus, faint events carry a limited amount of information. *Bottom panel:* $p^2_{\rm astro}$ as a function of squared SNR in this model. This quantity determines the average information contributed by each trigger.

The relative contributions to the information from loud and faint events can be different for other parameters, e.g. if the logarithmic terms in Eq. (4.38) preferentially select faint events. An example of this situation was demonstrated in Smith et al. (2020), who studied a putative cutoff in the distance distribution.

We choose the detection threshold (for actual triggers as well as for injections) as follows. We include all triggers from O1 and O2 that were found in Hanford–Livingston coincidence with a false-alarm rate (FAR)—within their chirp-mass bank, note that we searched in 5 such banks (Venumadhav et al., 2019; Roulet et al., 2019)—below one per three times the respective observing run. With this threshold choice, the summed $p^2_{\rm astro}(\lambda_0)$ of the 30 excluded triggers with lowest FAR is only 0.35, while for the events in Table 4.3.1 is 14.65. In order to include GW170817A consistently, we also include Livingston single-detector triggers from O2 that satisfy the cuts in templates and $\rho^2_{\rm L}$ described in Appendix 4.B (Zackay et al., 2019a). The exception is GWC170402, a Livingston single-detector trigger for which we do not have a satisfactory astrophysical model (Zackay et al., 2019a). We exclude GWC170402 from this analysis. From Table 4.3.1 we see that, with some dependence on the population model, including the events from the IAS catalog increases the accumulated $p^2_{\rm astro}$ by $\sim 47\%$, so as a crude estimation we might a priori expect uncertainties in the population parameters to shrink with a factor of order $\sim 1.47^{-1/2} = 0.82$ owing to this additional source of

information.

Note that Eq. (4.40) defines the Jeffreys prior for the astrophysical merger rate, $\pi(R \mid \lambda') \propto \sqrt{I(R)}$. In practice, $p_{\rm astro}$ depends on the rate only for a few near-threshold events and is otherwise very close to either 0 or 1, so we can approximate Eq. (4.40) by

$$I(R) \approx \frac{N_a(p_{\rm astro} \approx 1)}{R^2} \underset{\sim}{\propto} \frac{N_a(\lambda)}{R^2} = \frac{\overline{VT}(\lambda')}{R}. \tag{4.41}$$

Under this approximation we find

$$\pi(R \mid \lambda') \propto \sqrt{\overline{VT}(\lambda')/R}, \tag{4.42}$$

or $\pi(N_a) \propto 1/\sqrt{N_a}$, the Jeffreys prior for a single Poisson component. This differs from several other studies that adopt another prior, $\pi(R) \propto 1/\sqrt{R}$. In practice the difference is mild.

4.3 Astrophysical implications

In this section we report our results on the constraints on the BBH population parameters under various astrophysical models and compare their performances. To better visualize the effect of including the events from the IAS catalog, we repeat the analysis with a higher detection threshold (inverse false-alarm rate IFAR > 3000 observing runs in our pipeline) that restricts to events in the GWTC-1 catalog (Abbott et al., 2019a). GW170608 and GW170818 do not satisfy that cut so we include them ad hoc with $p_{\rm astro} = 1$, appropriately modifying the likelihood using Eq. (4.19) with both events.

We will explore a number of phenomenological models that probe the various measurable source parameters $m_{1\rm s}, q, \chi_{\rm eff}, D_L$. For convenience, we define a "default" distribution $\hat{f}$ that takes the following factorized form:

$$\hat{f}(m_{1\rm s}, q, \chi_{\rm eff}, D_L) = \hat{f}_{m_{1\rm s}}(m_{1\rm s})\, \hat{f}_q(q)\, \hat{f}_{\chi_{\rm eff}}(\chi_{\rm eff})\, \hat{f}_{D_L}(D_L), \tag{4.43}$$

	$p_{\rm astro}$	$\overline{p_{\rm astro}}$							
	λ_0	Default	A1	A2	A3	B	C	D	Combined
GW150914	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
GW170809	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
GW170104	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
GW170814	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
GW170729	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
GW170608	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
GW170823	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
GW151226	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
GW151012	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
GW170818	0.92	0.96	0.99	0.99	0.99	0.97	0.97	0.95	1.00
GW170304	1.00	0.99	1.00	1.00	1.00	0.99	1.00	0.99	1.00
GW170727	0.99	0.98	0.99	0.99	0.99	0.99	0.99	0.98	1.00
GW170121	0.98	0.99	0.98	0.98	0.97	0.99	0.99	0.98	0.99
GW170817A	0.75	0.27	0.14	0.20	0.20	0.07	0.43	0.30	0.01
GW170202	0.62	0.69	0.78	0.75	0.73	0.72	0.68	0.68	0.76
GW170403	0.62	0.53	0.16	0.12	0.08	0.50	0.61	0.53	0.11
GW170425	0.61	0.46	0.71	0.71	0.71	0.51	0.52	0.47	0.84
GW151216	0.51	0.47	0.00	0.01	0.00	0.55	0.52	0.47	0.00
170412B[2]	0.02	0.06	0.11	0.08	0.06	0.00	0.04	0.06	0.00

Table 4.3.1: Value of $p_{\rm astro}$ for the BBH events considered in this work under various astrophysical models. Events first reported in the GWTC-1 and IAS catalogs are respectively above and below the horizontal line. $p_{\rm astro}(\lambda_0)$ is the probability of astrophysical origin under the fiducial (unimportant) astrophysical model λ_0 described in §4.2.3. The remaining columns report the marginalized $\overline{p_{\rm astro}}$ under the population models considered in §4.3: the Default model (Eq. (4.43)) serves as a baseline from which Models A–E explore various departures as follows. A1: Gaussian in $\chi_{\rm eff}$; A2: tidally-locked progenitors with highly spinning remnants, or A3: with moderately spinning remnants; B: truncated power-law in the primary mass; C: power-law in the mass ratio; D: power-law in the redshift. The Combined model incorporates the maximum likelihood solutions of models A1, B and C (§4.3.6). $p_{\rm astro}(\lambda_0)$ approximately matches previously reported values from our pipeline (Venumadhav et al., 2019, 2020; Zackay et al., 2019a), except for GW170818 due to an improvement in the search algorithm, see Appendices 4.A and 4.B. We do not reproduce the results of (Galaudage et al., 2020, table II), see Appendix 4.C for a discussion.

with

$$\hat{f}_{m_{1s}} \propto m_{1s}^{-2.35}, \qquad 5\,\mathrm{M}_\odot < m_{1s} < 50\,\mathrm{M}_\odot \tag{4.44}$$

$$\hat{f}_q = \mathrm{U}(1/20, 1) \tag{4.45}$$

$$\hat{f}_{\chi_{\mathrm{eff}}} = \mathrm{U}(-1, 1) \tag{4.46}$$

$$\hat{f}_{D_L} = \frac{4\pi D_L^2}{(1+z)^4}\left(1 - \frac{D_L}{1+z}\frac{\mathrm{d}z}{\mathrm{d}D_L}\right), \tag{4.47}$$

and explore the effect of varying individual factors. $\mathrm{U}(a, b)$ denotes a uniform distribution between a and b. Note that $\hat{f}$ differs from $f(\lambda_0')$ (Eq. (4.24)) in that it has tighter lower and upper mass cutoffs, the merger rate is a free parameter, and the rate is set to be uniform in comoving volume–time (rather than luminosity volume and observer time) through the factor

$$\begin{aligned}\frac{\mathrm{d}t_c}{\mathrm{d}t}\frac{\mathrm{d}V_c}{\mathrm{d}V_L} &= \frac{\mathrm{d}t_c}{\mathrm{d}t}\frac{D_c^2}{D_L^2}\frac{\mathrm{d}D_c}{\mathrm{d}D_L} \\ &= \frac{1}{1+z}\cdot\frac{1}{(1+z)^2}\cdot\frac{1}{1+z}\left(1 - \frac{D_L}{1+z}\frac{\mathrm{d}z}{\mathrm{d}D_L}\right)\end{aligned} \tag{4.48}$$

in Eq. (4.47). $f(\lambda_0')$ is chosen to have support throughout all the sensitive parameter space, while $\hat{f}$ is meant to be a convenient reference point in the space of relevant population models and more closely comparable with the models explored in (Abbott et al., 2019b).

To obtain a posterior distribution for the population parameters, the likelihood Eq. (4.8) has to be multiplied by a prior. We take the priors to be flat except for the rate parameter, where we adopt a Jeffreys prior, Eq. (4.42).

For each of the models that we will consider below, in Table 4.3.1 we report for every event the astrophysical probability $\overline{p_{\mathrm{astro},i}}$ marginalized over uncertainties in the population model parameters:

$$\overline{p_{\mathrm{astro},i}} = \int \mathrm{d}\lambda\, P(\lambda \mid \{d\})\, p_{\mathrm{astro},i}(\lambda). \tag{4.49}$$

To evaluate Eq. (4.49), we obtain $p_{\mathrm{astro},i}(\lambda)$ from Eqs. (4.7) and (4.12) as

$$p_{\mathrm{astro},i}(\lambda) = \frac{\dfrac{R}{R_0} w_i(\lambda'; \lambda_0')\, p_{\mathrm{astro},i}(\lambda_0)}{1 + \left(\dfrac{R}{R_0} w_i(\lambda'; \lambda_0') - 1\right) p_{\mathrm{astro},i}(\lambda_0)} \tag{4.50}$$

and marginalize over λ by quadrature.

4.3.1 Spin distribution

Model A1: Gaussian $\chi_{\rm eff}$ distribution

We first consider a distribution that is Gaussian in $\chi_{\rm eff}$

$$f_{\chi_{\rm eff}}(\chi_{\rm eff} \mid \overline{\chi_{\rm eff}}, \sigma_{\chi_{\rm eff}}) = \mathcal{N}_{\rm t}(\chi_{\rm eff} \mid \overline{\chi_{\rm eff}}, \sigma_{\chi_{\rm eff}}) \tag{4.51}$$

and follows the default $\hat{f}$ (4.43) in the other parameters. $\mathcal{N}_{\rm t}(x \mid \mu, \sigma)$ denotes the normal distribution with mean μ and dispersion σ truncated at ± 1. With this simple model we can explore the symmetry of the $\chi_{\rm eff}$ distribution, i.e. whether there is a tendency for alignment between the spins and orbit or not. Dynamical formation models predict the spins of the black holes to be randomly oriented, thus symmetrically distributed about $\chi_{\rm eff} = 0$, while in the isolated binary evolution scenario a tendency for alignment might be expected (note that while spin-dependent selection effects can bias the detected distribution towards positive effective spins (Ng et al., 2018; Roulet and Zaldarriaga, 2019), f describes the underlying astrophysical distribution). This model also probes the width of the distribution, which is especially interesting in light of predictions that natal BH spins might be very small, barring tidal torques to the stellar progenitors (Fuller and Ma, 2019).

The results are shown in Fig. 4.3.1. In agreement with previous studies (Farr et al., 2018; Wysocki et al., 2018, 2019; Roulet and Zaldarriaga, 2019; Abbott et al., 2019b; Miller et al., 2020), the distribution is consistent with $\overline{\chi_{\rm eff}} = 0$ and shows no statistically significant preference for positive $\chi_{\rm eff}$, which remains consistent with dynamical formation scenarios. The width of the distribution is measured to be $\sigma_{\chi_{\rm eff}} \approx 0.13^{+0.12}_{-0.07}$ (median and 90% c.l.), disfavoring values close to 0 (cf. Miller et al., 2020). Including the events from the IAS catalog yields consistent results. The constraints become tighter for the rate, and broader for the $\chi_{\rm eff}$ mean and dispersion. We verified that GW170121, a confident detection with support for negative $\chi_{\rm eff}$ (Venumadhav et al., 2020), rules out the end of the distribution compatible with GWTC-1 with higher $\overline{\chi_{\rm eff}}$ and smaller $\sigma_{\chi_{\rm eff}}$, which may drive the change in these constraints. In the rest of this section, we will explore whether the spread of the $\chi_{\rm eff}$ distribution can be explained by tidal torques.

Model A2: tidally-locked stellar progenitors

We now study a model that considers the effects of tides in the BBH progenitor system. For field binaries, the typical aftermath of the common-envelope phase is a black hole in a tight orbit with a stripped star. Depending on the binary separation, the star may be subject to strong tides. If

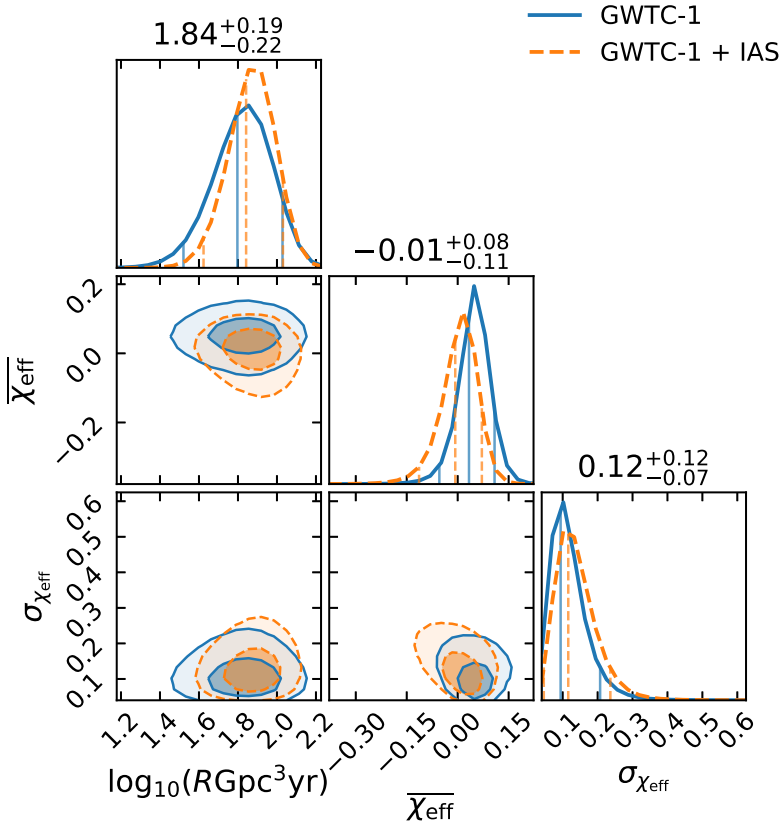


Figure 4.3.1: Model A1: Gaussian $\chi_{\rm eff}$ distribution. Solid blue curves show the constraints on the model derived from only the GWTC-1 catalog. Dashed orange curves show the constraints derived from the GWTC-1 and IAS catalogs combined. Two-dimensional contours enclose 50% and 90% of the distribution. Vertical lines show the median and 90% symmetric interval of the one-dimensional posteriors, also reported numerically for the GWTC-1 + IAS analysis.

the separation is small, corresponding to merger times $\lesssim 10^8\,\mathrm{yr}$, the timescale for tidal locking is shorter than the star lifetime and thus it can tidally lock to the orbit. In this case the second-formed black hole would have a high, aligned spin. The tidal-locking timescale depends strongly on the separation, so that for greater separations tides quickly become negligible. In addition, the maximum separation allowed for a circular binary to merge within the age of the Universe is comparable to this distance scale, so two sub-populations with comparable abundances might be expected (Kushnir et al., 2016; Hotokezaka and Piran, 2017; Zaldarriaga et al., 2017). It has been pointed out that with different wind and tide models, less extreme distributions may result, with the possibility of having intermediate spins even after tidal locking (Qin et al., 2018; Bavera et al., 2020). We will first probe the more extreme model which is easier to constrain, keeping in mind that the bounds we obtain apply to the fraction of black holes with maximal spin, and later explore the consequences of a milder spin distribution.

Following Ref. (Roulet and Zaldarriaga, 2019), we implement a model of this scenario as follows. We consider that component black holes have $\chi_z \sim \mathcal{N}_\mathrm{t}(0, \sigma_\chi)$ in the absence of tidal effects, and a fraction ζ of the secondary (lighter) BHs have $\chi_z = 1$ due to a tidally locked progenitor. This distribution is very different from both the injection distribution and the parameter estimation prior in the space of component spins. In order to have well-behaved reweighting of samples (see §4.2.3), in practice we implement it under the approximation that the strain waveform depends on the spins only through χ_eff. We therefore use

$$\begin{aligned} f_{\chi_\mathrm{eff}}(\chi_\mathrm{eff} \mid q, \zeta, \sigma_\chi) &= \iint \mathrm{d}\chi_{1z}\mathrm{d}\chi_{2z} f(\chi_{1z}, \chi_{2z} \mid \zeta, \sigma_\chi)\delta\Big(\chi_\mathrm{eff} - \frac{\chi_{1z} + q\chi_{2z}}{1+q}\Big) \\ &= (1-\zeta)\,\mathcal{N}_\mathrm{t}(\chi_\mathrm{eff} \mid \mu_0, \sigma_0) + \zeta\,\mathcal{N}_\mathrm{t}(\chi_\mathrm{eff} \mid \mu_1, \sigma_1), \end{aligned} \tag{4.52}$$

where the subscript 0 represents the case in which no tidal locking occurred and 1 the case where the progenitor of the secondary was locked:

$$\begin{aligned} \mu_0 &= 0 \qquad & \sigma_0 &= \frac{\sqrt{1+q^2}}{1+q}\sigma_\chi \\ \mu_1 &= \frac{q}{1+q} & \sigma_1 &= \frac{\sigma_\chi}{1+q}. \end{aligned} \tag{4.53}$$

We show our constraints in the top panel of Fig. 4.3.2. As in Ref. (Roulet and Zaldarriaga, 2019), we find that the fraction of locked systems ζ is consistent with 0. We bind it to $\zeta < 0.2$ at 90% confidence. Interestingly, the extreme version of this model $\sigma_\chi \approx 0$, where black holes are born with nearly zero spins except for tidal effects (Fuller and Ma, 2019), is in mild tension. Even when

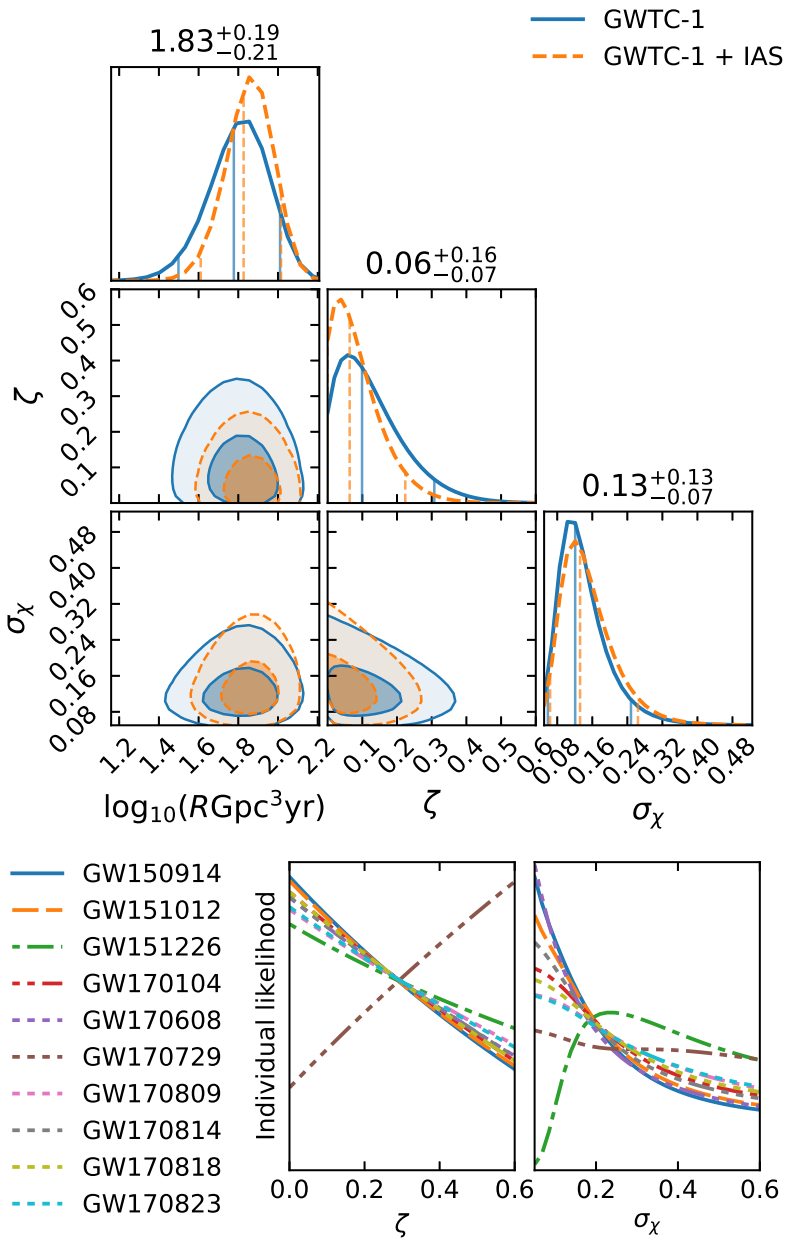


Figure 4.3.2: Model A2: a fraction ζ of the secondary black holes have tidally locked progenitors ($\chi_{2z} = 1$) and the remaining $(1-\zeta)$ fraction follow the Gaussian distribution $\chi^z_{1,2} \sim \mathcal{N}_\mathrm{t}(0, \sigma_\chi)$. *Top panel:* Population parameter constraints. *Bottom panels:* Individual marginalized likelihood for each of the events in the GWTC-1 catalog.

restricted to the GWTC-1 catalog, this conclusion is in disagreement with (Roulet and Zaldarriaga, 2019). The immediate cause of the difference is that in this work we found shorter tails in the mass-ratio distribution of some events, and these tails can affect how consistent an event is with the maximally-spinning-secondary hypothesis.[3] The main differences between the parameter estimation methods of the two analyses, which could explain the disagreement, are that in this work we apply an exact treatment of the detector locations and orientations instead of the approximations made in (Roulet and Zaldarriaga, 2019) and we have a more careful estimation of the noise PSD (Zackay et al., 2019c; Huang et al., 2020), thus the new results are preferred. In Fig. 4.3.3 we show the events' posteriors in the q–$\chi_{\rm eff}$ plane, and the curve $\chi_{\rm eff} = q/(1+q)$ corresponding to $\chi_{1z} = 0$, $\chi_{2z} = 1$ as representative of the tidally-locked progenitor scenario. In particular GW151226, whose effective spin is positive and well-measured, is only marginally consistent with having a χ_{2z} as high as 1 and prefers lower values. In the bottom panels of Fig. 4.3.2 we show this situation in further detail. For the events in the GWTC-1 catalog, we plot the individual marginalized likelihoods (i.e. the terms inside the product in Eq. (4.9)) for ζ and σ_χ. We see that the low-σ_χ solution is disfavored by GW151226. Additionally, GW170729 is more consistent with having $\chi_{2z} = 1$ than with coming from a low spin distribution, thus it pushes the locked fraction upwards.

To keep the number of parameters small, in Model A2 we have implicitly assumed that the second-formed black hole, whose progenitor can be subject to strong tides, is the lightest. While lighter stars have longer lifetimes in isolation, for binary evolution the mass ratio may be reversed by mass transfer episodes (Gerosa et al., 2013; Steinle and Kesden, 2020). Moreover, depending on the detailed ordering of mass transfer and core collapses both stars may be subject to strong tides (Steinle and Kesden, 2020). We explored the impact of this simplification using a model where either BH has small or high-aligned spin independently with some probability, when we do this we obtain similar constraints on the total fraction of tidally-locked progenitors and natal spin dispersion as with Model A2.

Model A3: tidally-torqued progenitors with moderately spinning remnants

Is the tension with very low natal BH spins robust to a less extreme model for the spin of remnant of a tidally torqued star? Motivated by (Qin et al., 2018; Bavera et al., 2020) we now study a version of the model where these black holes have spin aligned with the orbit, but with a broad distribution, which we take to be $P(\chi_{2z}) = \mathrm{U}(0,1)$ for simplicity. We assume that otherwise all black holes have

[3] Parameter estimation samples are available at `github.com/jroulet/O2_samples`.

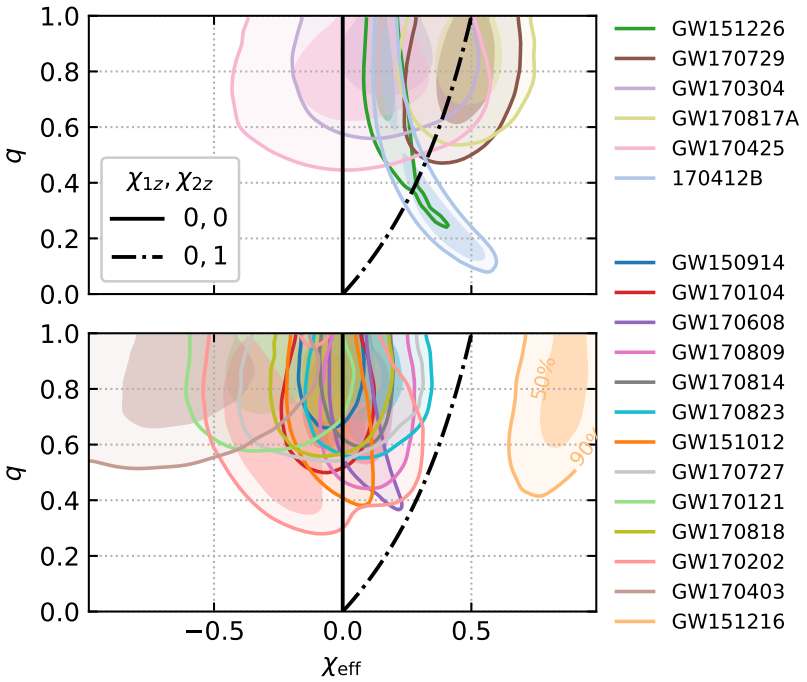


Figure 4.3.3: Posterior on the q–$\chi_{\rm eff}$ plane for the events considered in this work that are consistent (top) or inconsistent (bottom) with having a non-spinning primary and a maximally spinning, aligned secondary at the 90% confidence level.

very low spin (Fuller and Ma, 2019). We parameterize the $\chi_{\rm eff}$ distribution as

$$f_{\chi_{\rm eff}}(\chi_{\rm eff} \mid q; \zeta) = \zeta\, \mathrm{U}(\chi_{\rm eff} \mid 0, q/(1+q)) + (1-\zeta)\mathcal{N}_{\rm t}(\chi_{\rm eff} \mid 0, 0.05). \tag{4.54}$$

For the subpopulation where tides did not play an important role, we have set a nonzero width in $\chi_{\rm eff}$ comparable to measurement uncertainties, and made the approximation that the likelihood depends on the spins through $\chi_{\rm eff}$ only, in order to avoid pathological reweighting. We show the constraints on Fig. 4.3.4. Interestingly, under these assumptions we find that a fraction $\zeta = 0.22^{+0.29}_{-0.19}$ of the BBHs needs to have been subject to strong tidal effects in order to explain the events that are measured to have positive $\chi_{\rm eff}$. We have verified that if we allow for a further spread σ_χ in the natal BH spin distribution, this spread is well consistent with 0, unlike for Model A2.

In conclusion, the data is consistent with either a natal distribution of spins with nonzero dispersion $\sigma_\chi \sim 0.1$, or a model where the remnant may have a moderate spin even after tidal torquing. In the first case, the result holds even if we allow a fraction of tidally-locked progenitors (the inferred fraction is consistent with zero); in the latter, which better aligns with the predictions of Refs. (Qin et al., 2018; Fuller and Ma, 2019; Bavera et al., 2020), a fraction $\zeta \sim 0.2$ of tidally torqued events is favored. Later, in §4.3.6, we will show results for the relative evidence for these families of models

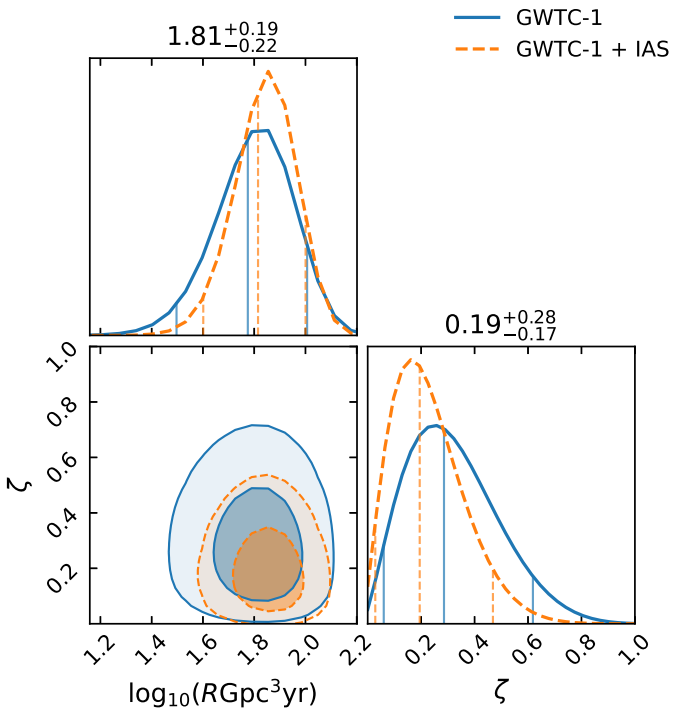


Figure 4.3.4: Model A3, a modified version of Model A2 where the remnant of the tidally-locked progenitor has χ_{2z} distributed uniformly in $[0, 1]$ for a fraction ζ of the mergers, and otherwise black holes have very low spins.

and find some preference for Model A3 over A1 and A2. Again, we find that including the events from the IAS catalog improves the constraints on these population models.

4.3.2 Model B: truncated power-law in primary mass

Next we consider a model where the source-frame mass of the primary black hole follows a truncated power-law distribution

$$f_{m_{1s}}(m_{1s} \mid \alpha, m_{\min}, m_{\max}) \propto m_{1s}^{-\alpha};\ m_{\min} < m_{1s} < m_{\max} \tag{4.55}$$

$$f_q(q \mid m_{1s}, m_{\min}) = \mathrm{U}(m_{\min}/m_{1s}, 1). \tag{4.56}$$

Eq. (4.56) enforces that the secondary mass also satisfies the lower cutoff. This model was first studied by Ref. (Fishbach and Holz, 2017) (see also (Roulet and Zaldarriaga, 2019; Abbott et al., 2019b)) and is motivated by the prediction of a gap in the stellar black hole mass function due to the pair instability supernova and pulsational pair instability supernova processes (Fowler and Hoyle, 1964; Barkat et al., 1967; Bond et al., 1984; Heger et al., 2003).

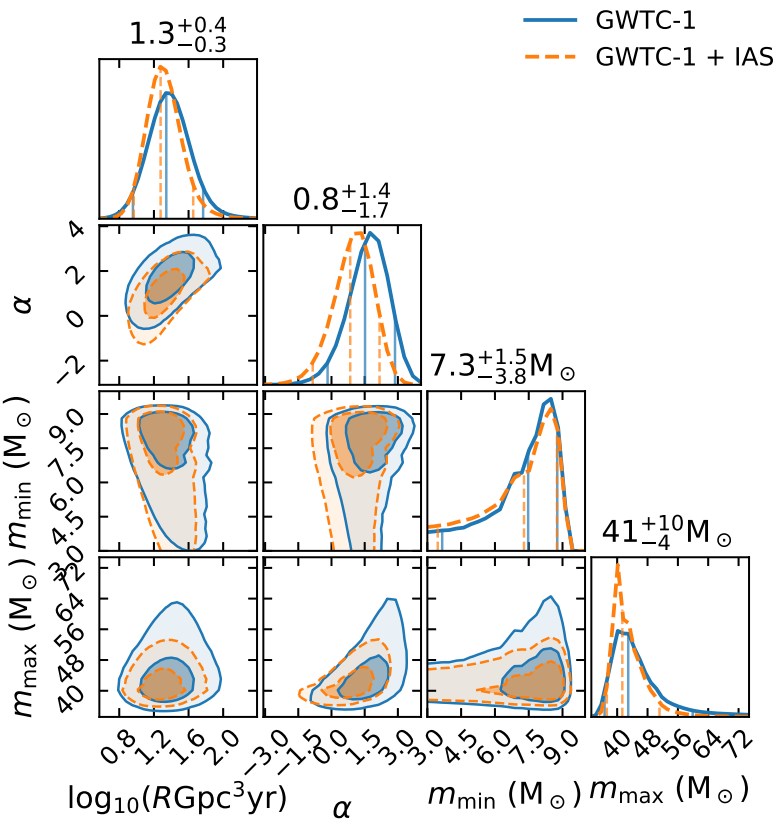


Figure 4.3.5: Model B: a power-law distribution for the primary black hole mass truncated at the minimum and maximum masses.

We show our constraints in Fig. 4.3.5. These might be approximately compared to the models explored in (Abbott et al., 2019b, figure 2). When restricting to the events in GWTC-1 we obtain similar results to Ref. (Abbott et al., 2019b), except we recover a merger rate that is lower by ~ 0.4 decades, which is smaller than the current statistical uncertainty. Interestingly, we obtain a tighter constraint $m_{\rm max} < 51\,{\rm M}_\odot$ at 90% confidence by including the events from the IAS catalog. This is accompanied by a slight shift in the allowed power-law index towards shallower slopes. A putative lower cutoff $m_{\rm min}$ in the mass function is harder to detect since low-mass mergers are intrinsically fainter. Thus, similar to Abbott et al. (2019b), we can only put an upper bound on $m_{\rm min}$ given by the lightest confident merger considered, GW170608.

We comment that the O3 event GW190521 has a primary black hole mass $85^{+21}_{-14}\,{\rm M}_\odot$ (Abbott et al., 2020c), in tension with the constraints in Fig. 4.3.5. This suggests that the truncated power-law model may not be a good description of the tails of the distribution once a larger number of events is included, and that parametrizations with more complexity might be needed in future analyses.

4.3.3 Model C: power-law in the mass ratio

Now we study a model where the mass ratio follows a power law distribution (Roulet and Zaldarriaga, 2019; Fishbach and Holz, 2020)

$$f_q(q \mid \overline{q}) \propto q^\beta \tag{4.57}$$

with $\beta = (2\,\overline{q}-1)/(1-\overline{q})$ so that the distribution has a mean $\overline{q}$. We show our constraints in Fig. 4.3.6. In line with previous results (Roulet and Zaldarriaga, 2019; Abbott et al., 2019b; Fishbach and Holz, 2020), we find that distributions leaned towards equal mass ratios are favored. Including the events from the IAS catalog enables a more precise measurement of both the rate and mean mass-ratio. We find a mild quantitative difference in the $\overline{q}$ distribution with (Roulet and Zaldarriaga, 2019, figure 9), which we verified is due to the difference in the underlying mass distribution models (truncated power-law in primary mass per Eq. (4.44) vs. flat in chirp mass); the one used here is strongly favored in terms of model selection by a difference in log-likelihood $\Delta \max_\lambda \ln P(\{d_i\}, N_{\rm trig} \mid \lambda) = 24$.

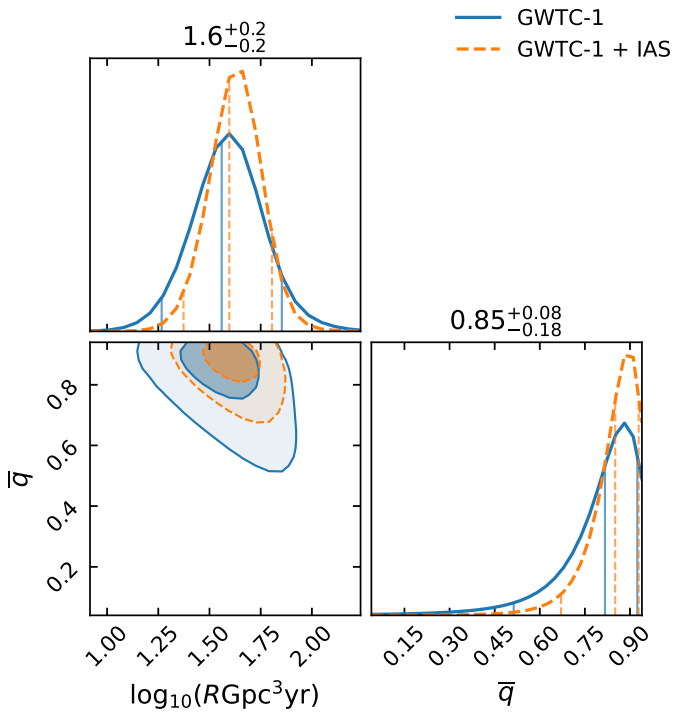


Figure 4.3.6: Model C: a power-law in the mass ratio with mean $\overline{q}$, Eq. (4.57).

4.3.4 Model D: power-law in the redshift

Finally, we consider a model where the comoving merger rate follows a power-law in the redshift with index λ_z (Fishbach et al., 2018)

$$f_{D_L}(D_L \mid \lambda_z) = \hat{f}_{D_L}(D_L)(1+z)^{\lambda_z}, \tag{4.58}$$

so that $\lambda_z = 0$ corresponds to a constant merger rate per unit comoving volume (see Eq. (4.47)). We show the constraints on λ_z in Fig. 4.3.7: it is poorly constrained and is consistent with a constant merger rate per unit comoving volume. Adding the new events from the IAS catalog improves the constraints both on the rate and on λ_z.

4.3.5 Merger rate

As seen from Figs. 4.3.1–4.3.7, the measurements of the local merger rate R using each of the models we studied are consistent within uncertainties, however, they have broad distributions and may have large correlations with some of the population parameters. This happens because the rate parameter is measured through the combination $N_a(\lambda) = R \cdot \overline{VT}(\lambda')$, and the population-averaged sensitive-

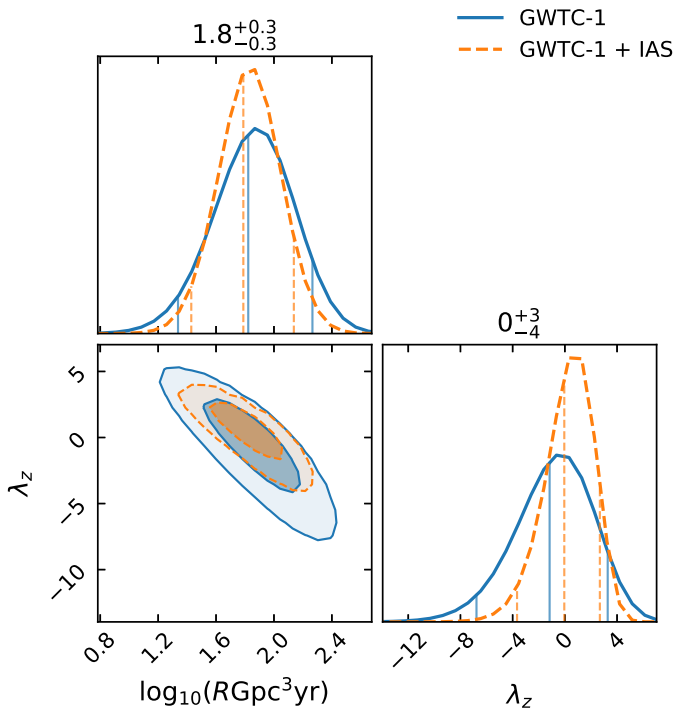


Figure 4.3.7: Model D: a power-law in the redshift evolution of the merger rate with exponent λ_z, Eq. (4.58). R is the local rate.

volume $\overline{VT}(\lambda')$ can exhibit a large dependence on the population shape. In other words, most of the information comes from the region in parameter space where most events lie. This region depends on the interplay between the astrophysical population and the detector selection function; the rate within this region should be relatively well constrained compared to the overall rate.

The most extreme example is Model D (power-law in the redshift, Fig. 4.3.7), where the rate exhibits a strong correlation with the exponent λ_z. In this case, the merger rate at some intermediate redshift $z \sim 0.2$ is much better constrained than the local rate R, owing to the larger phase space that the detectors are sensitive to. To various extents, a similar effect occurs with some shape parameters in the other models studied in this section. In Model B (power-law in the primary mass, Fig. 4.3.5), for steep power-laws the rate can be dominated by the low-mass end of the distribution, which is poorly constrained because the sensitive volume to these signals is smaller.

A simple prescription to get a robust constraint that can inform theoretical models is to measure the rate of mergers within the part of parameter space where most events were observed. In Fig. 4.3.8 we plot the posterior distribution for the restricted rate of mergers $R_{\mathrm{restricted}}$, which we define as the merger rate of events with $20\,\mathrm{M}_\odot < m_{1\mathrm{s}} < 30\,\mathrm{M}_\odot$, $q > 0.5$, evaluated at redshift $z = 0.2$. We show this quantity for the default model (Eq. (4.43)), Models A1–D (§A1–4.3.4), and a 'Combined' model

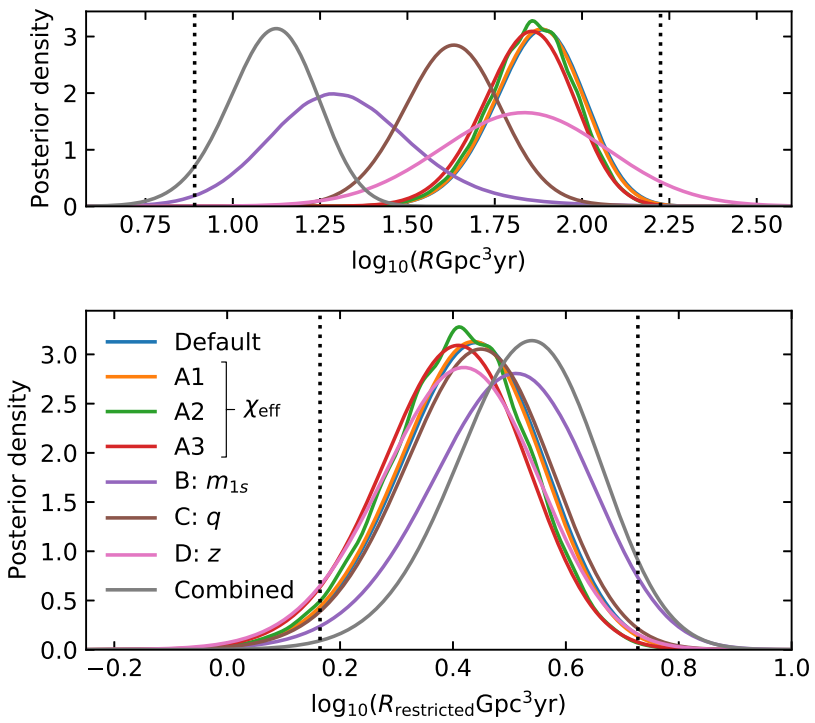


Figure 4.3.8: Posterior for the local rate of mergers (top) or the comoving rate restricted to $20\,\mathrm{M}_\odot < m_{1\mathrm{s}} < 30\,\mathrm{M}_\odot$, $q > 0.5$ and evaluated at redshift $z = 0.2$ (bottom), under the various models we studied (§A1–4.3.4, annotated with the variable each of them is intended to explore). We add a model that combines the maximum likelihood solutions of models A1, B and C. The restricted rate is better measured, and in better agreement across models, than the total local rate R. Dotted lines indicate the minimum over models of the 5% quantile, and the maximum of the 95% quantiles.

that combines the maximum likelihood solutions of models A1, B and C, namely the product of $f_{\chi_{\mathrm{eff}}}$ from Eq. (4.51) with $\overline{\chi_{\mathrm{eff}}} = 0, \sigma_{\chi_{\mathrm{eff}}} = 0.1$; $f_{m_{1\mathrm{s}}}$ from Eq. (4.55) with $\alpha = 1, m_{\mathrm{min}} = 8.5\,\mathrm{M}_\odot, m_{\mathrm{max}} = 40\,\mathrm{M}_\odot$; f_q from Eq. (4.57) with $\overline{q} = 0.88$; and $\hat{f}_{D_L}$ from Eq. (4.47). Note that the primary mass is distributed with a truncated power-law of index -2.35 in all except the B and Combined models, and for the Combined model it is flat-in-log, which makes Fig. 4.3.8 largely comparable to (Abbott et al., 2019a, fig. 12); this may also be the main driver of the residual discrepancies between restricted rates across models in Fig. 4.3.8.

We find that the restricted rate is much better measured than the total local merger rate and that all models largely agree on its value. Taking the union of the symmetric 90% quantiles, we obtain that the rate lies in the range 1.5–5.3 $\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}$, which amounts to a factor 3.6 uncertainty in the restricted rate as opposed to 21 in the absolute rate, within the models we explored. Thus, the restricted rate is well suited to put population models to a more stringent test.

Model	$\Delta \max \ln L$	$\Delta \ln Z$
Default	0	0
A1 ($\chi_{\rm eff}$)	$14.6^{+1.3}_{-1.1}$	$10.3^{+1.1}_{-0.8}$
A2 ($\chi_{\rm eff}$)	$14.9^{+1.3}_{-0.9}$	$12.0^{+1.1}_{-0.8}$
A3 ($\chi_{\rm eff}$)	$14.6^{+1.5}_{-1.2}$	$13.4^{+1.4}_{-1.1}$
B: m_{1s}	$3.8^{+0.6}_{-0.5}$	$-2.1^{+0.5}_{-0.3}$
C: q	$4.6^{+0.6}_{-0.6}$	$2.8^{+0.5}_{-0.4}$
D: z	$0.07^{+0.19}_{-0.06}$	$-2.84^{+0.29}_{-0.12}$
Combined	22^{+3}_{-3}	22^{+2}_{-2}

Table 4.3.2: Scores for the population models we study (§4.3.1–4.3.4, annotated with the variable each of them is intended to explore) relative to the Default model, Eq. (4.43). We add a model that combines the maximum likelihood solutions of models A1, B and C. We report the maximum likelihood and the Bayesian evidence as complementary indicating scores. Error-bars indicate 90% confidence levels on the uncertainties from the Monte Carlo method employed.

4.3.6 Comparison between population models

We conclude this section by comparing the performances of the models we studied. In Table 4.3.2 we provide the maximum likelihood and the Bayesian evidence achieved by models A1–D and Combined discussed above, relative to the default model Eq. (4.43). The maximum likelihood has the advantage of being independent of the arbitrary prior choices for the population parameters (in particular, their ranges), on the other hand it does not penalize models with more degrees of freedom. The ordering and qualitative conclusions are found to be largely similar using either metric. Our method involves multiple Monte Carlo integrations (numerator and denominator of Eq. (4.17), and Eq. (4.18)), which can introduce stochastic errors. We estimate these with the bootstrap method, by computing the scores 100 times using samples taken with replacement from the original sets.

We find that among Models A1–D, those that perform best are A1 (Gaussian in $\chi_{\rm eff}$), A2 and A3 (tidally locked progenitors, with $\chi_{2z} = 1$ or $\chi_{2z} \sim \mathrm{U}(0,1)$ respectively), which vary the spin distribution away from the flat-in-$\chi_{\rm eff}$ one the default model uses. This is a clear indication that the average effective spins of the population are lower. The three models achieve similar likelihoods, Model A3 has a somewhat higher Bayesian evidence, which can be related to the fact that it has one less parameter. Note that these models cannot accommodate the high effective spin of GW151216, which requires both spins to be high and aligned; its $\overline{p_{\rm astro}}$ is suppressed as a result (Table 4.3.1). To a lesser extent, a similar effect holds for GW170817A and GW170403.

The Combined model, which is defined from the best likelihood solutions of Models A1, B and C, indeed outperforms its individual components and, as expected, the likelihood ratio to the default model is approximately the product of the component likelihood ratios to the default model. This vindicates our approach of individually varying the components of the population model.

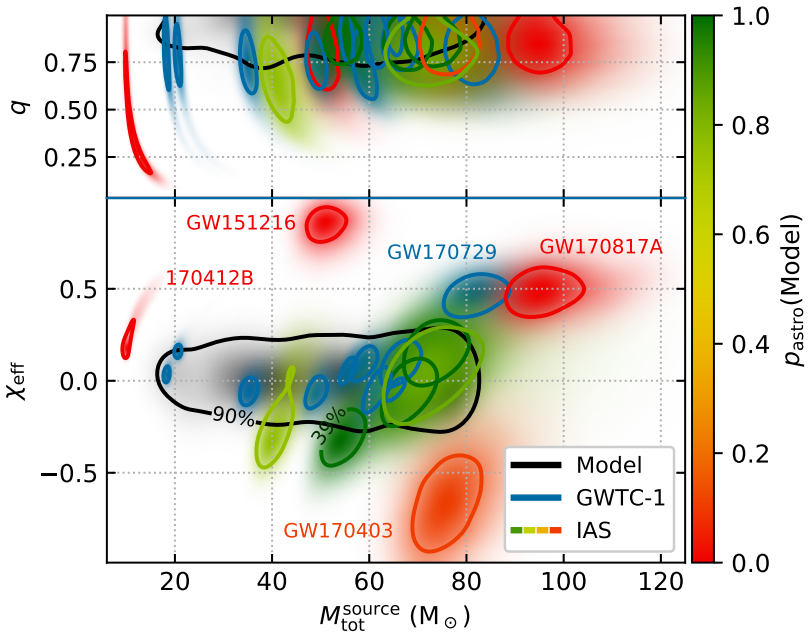


Figure 4.3.9: Events considered in this work as a function of total source-frame mass, mass-ratio and effective spin. Underlaid in black is the prediction of a model that combines the maximum likelihood solutions for the effective spin (Model A1), primary mass (Model B) and mass-ratio (Model C) distributions, including selection effects and without measurement uncertainty. The events' posteriors are color-coded according to their values of $p_{\rm astro}$ under the same model (these are different from the values reported by the search pipeline). Note however that these posteriors are computed using a prior that differs from the model. Model outliers are labeled.

In Fig. 4.3.9 we plot the BBH events from O1 and O2 considered in this work, showing posteriors in the space of total source-frame mass versus mass-ratio and effective spin. 1σ contours (enclosing $1 - e^{-1/2} \approx 0.39$ of the distribution) are drawn in blue for events in the GWTC-1 catalog or color-coded by $p_{\rm astro}$ for events in the IAS catalog. The broad parameter estimation prior $\pi(\theta)$, defined in §4.2.3, is used for event contours in order to make the resulting posteriors trace more closely the single-event likelihoods from Eq. (4.3). Shown in black is the expected distribution of detectable sources under the Combined model, with a 90% contour. It is obtained by reweighting found injections with this model. The $p_{\rm astro}$ values shown in the color scale of Fig. 4.3.9 correspond to this model. Note that this is just one example out of the set of models consistent with the data, and others may exhibit somewhat different behavior.

This simple model appears to explain the observations reasonably well, albeit with some outliers. The confident event GW170729 (Abbott et al., 2019a; Venumadhav et al., 2020) is a mild outlier within 2σ of the 90% contour, note however that under this population model its posterior would shift towards lower mass and $\chi_{\rm eff}$ (Fishbach et al., 2020). The other outliers (GW170817A, GW170403, GW151216 and 170412B) have less detection significance, so their $p_{\rm astro}(\lambda)$ can be suppressed without

too large a penalty to the model likelihood. As our knowledge of the astrophysical distribution improves, the $p_{\rm astro}$ of these events might be revised. A potential limitation of the models studied in this work is the assumption that the distribution is factorizable in the variables $m_{1s}, q, \chi_{\rm eff}, D_L$ (Models A2 and A3 do have a correlation between q and $\chi_{\rm eff}$, and Model B between m_{1s} and q). Indeed, Fig. 4.3.9 might hint that models in which the $\chi_{\rm eff}$ mean or variance depends on the mass (Safarzadeh et al., 2020) or distance (Hotokezaka and Piran, 2017) could perhaps better accommodate some of these outliers. Looking forward, as the catalog of events grows there might be a need for increasingly complex models.

We comment that 170412B is consistent with having a secondary mass similar to that of the O3 event GW190814 (Abbott et al., 2020d), although its $\chi_{\rm eff}$ would be higher especially if such a small mass-ratio is imposed.

4.4 Conclusions

In this work, we presented a general framework to include marginal GW events when characterizing the astrophysical population of BBH mergers. Similar to other proposed methods, ours requires that we characterize the sensitivity of the search pipeline. We achieved this using software injections, which we generated using a fiducial distribution and reweighted as needed. We generated the parameters of the injected events from a reference population model weighted by an approximate detector sensitivity function; this simple prescription achieved a good balance between accepted and rejected injections. We implemented an injection campaign using the above strategy, and empirically measured the sensitive volume–time of our search to be $\simeq 0.17\,{\rm Gpc}^3\,{\rm yr}$, for the population model in Eq. (4.43) and with a FAR threshold of one per all our O1 and O2 BBH searches combined (Appendix 4.A.3). In its own right, quantifying the sensitivity of our pipeline solidifies the evidence in favor of the detections that we previously reported.

We demonstrated our method by using events reported in the IAS catalog, in addition to the ones originally reported by the LVC, to characterize the BBH population. In particular, we studied various phenomenological population models that explored the spin, mass, mass-ratio and redshift distributions. For models that have been previously explored, our results are broadly consistent with previous studies, with reduced uncertainties due to the extra information that the additional events contribute. We quantified the information gain from including these additional events, e.g. for the astrophysical rate, it scales with their summed $p^2_{\rm astro}$ and amounts to a $\sim 47\%$ increase.

In models where the effective spin parameter, $\chi_{\rm eff}$, of all mergers is drawn from a normal dis-

tribution, we do not find any statistically significant deviations from $\overline{\chi_{\rm eff}} = 0$, and infer a typical spread of $\sigma_{\chi_{\rm eff}} \sim 0.1$. If we allow a fraction ζ of the secondary black holes to have aligned spins due to tidal effects on their progenitors, the conclusions depend on how efficient the tides can be. If the tides, when operative, are strong enough that the secondary BBH ends up with maximal spin, the fraction ζ is consistent with zero and bounded to $\zeta < 0.2$ (the spread $\sigma_{\chi_{\rm eff}}$ of the mergers without tides is similar to the previous case). If the tides are weak enough or the details of the collapse allow for an aligned, but not necessarily maximally-spinning secondary, the data can be explained by a fraction $\zeta \sim 0.2$ of BBHs with tidally torqued progenitors, with the rest having very low natal spins. Future data might be able to distinguish between these two scenarios; it is intriguing that two of the BBH mergers reported so far in the O3 run have non-zero and positive values of $\chi_{\rm eff}$ (Abbott et al., 2020b,d).

The mass, mass-ratio, and redshift distributions are consistent with previous work: if the masses are drawn from a truncated power-law distribution, we bound the upper cutoff in the primary mass to $m_{\rm max} < 51\,{\rm M}_\odot$ at 90% confidence. The data favor a mass-ratio distribution that leans towards the equal-mass case, $\overline{q} > 0.67$, and a redshift distribution that is consistent with uniform in comoving volume.

We additionally argued that the merger rate is better measured if restricted to the region of parameter space where most events are found. We find that the merger rate restricted to BBHs with a primary mass between 20–30 $\rm M_\odot$, mass ratio $q > 0.5$, and at $z \sim 0.2$, is 1.5–5.3 $\rm Gpc^{-3}yr^{-1}$ (90% c.l.). Unlike for the total local merger rate, this constraint is model independent and a factor of ~ 3 tighter, and thus well-suited for testing progenitor models.

Apart from the results on the population models for the data we included, we foresee that the methodology presented here will continue to prove useful as future data releases will generically include marginal detections. We have emphasized the dependence of $p_{\rm astro}$ on the astrophysical model and the search pipeline used. An intermediate step in our method is to compute the $p_{\rm astro}$ of the triggers of interest for a specified reference source population model, which we make as permissive as possible to facilitate reweighting. Looking forward, this can be a convenient convention for reporting values of $p_{\rm astro}$, especially for marginal triggers whose final interpretation may depend on the population model.

Appendix

4.A Sensitivity of the search pipeline

In this appendix we report results of our injection campaign and characterize the sensitivity of our primary search for BBH mergers (Venumadhav et al., 2019, 2020).

4.A.1 Injection campaigns

We make 50 000 software injections in each of the O1 and O2 observing runs, at random times without regard to the duty cycle of the detectors, with source parameters distributed according to Eq. (4.25). We then run all stages of our search as described in Refs. (Venumadhav et al., 2019, 2020), except for the following two modifications.

First, for the injection campaigns we disable the initial stage of noise transient (glitch) rejection and inpainting (Venumadhav et al., 2019; Zackay et al., 2019c), which would otherwise greatly increase the computational cost. Instead, we keep track of the locations where glitches were identified in the original search. The pipeline does not record triggers within 1 s of an identified glitch for templates shorter than 10 s, so we treat those times as invalid for observation.

Second, we implement an improved version of the coherent score—this is the piece of the detection statistic that accounts for signal coherence across detectors (Nitz et al., 2017; Venumadhav et al., 2019), and depends on the difference in the arrival times, as well as the relative phase between the detectors. Our implementation of the coherent score uses the best measured values of these parameters from the data, and accounts for the requisite amounts of measurement noise in each of these values. However, in previous work, we had neglected the effect of the correlation between the measurements of the arrival times and phases that are input to the coherent score; we found that this ultimately caused us to assign a high FAR to the LVC event GW170818 in our coincidence search (using only Hanford and Livingston data;[4] the FAR was however not biased since this effect

[4]Our search pipeline is so far restricted to LIGO Hanford and Livingston data, on which all our significance

Run	Event	GPS time	IFAR (run)
O1	GW150914	1126259462.41	$> 20\,000$
	GW151012	1128678900.43	$> 20\,000$
	GW151226	1135136350.59	$> 20\,000$
	GW151216	1134293073.16	26.70
O2	GW170823	1187529256.50	$> 20\,000$
	GW170809	1186302519.74	$> 20\,000$
	GW170729	1185389807.31	$> 20\,000$
	GW170814	1186741861.52	$> 20\,000$
	GW170104	1167559936.58	$> 20\,000$
	GW170727	1185152688.02	256.41
	GW170121	1169069154.57	185.19
	GW170304	1172680691.36	78.74
	GW170818	1187058327.08	30.40
	170412B[5]	1175991666.07	6.51
	GW170403	1175295989.22	6.25
	GW170425	1177134832.18	5.30
	GW170202	1170079035.72	4.19

Table 4.A.1: Inverse false-alarm rate for the events considered in this work assigned by our primary search (Venumadhav et al., 2019, 2020) with an improved detection statistic. Note that GW170817A was found in a different, targeted search (Zackay et al., 2019a). The IFAR reported here is referred to the individual observing run and template bank each event was found in, without penalizing for the fact that we searched two observing runs and five template banks, see §4.A.3. These are all events we found with Hanford–Livingston coincident triggers with an IFAR > 3 runs in their bank.

impacted the timeslides as well). The new version we use in this paper accounts for the correlations in the measurements, and hence should be closer to optimality (in the sense of being closer to the likelihood ratio test). When we apply this improvement to the coincidence search itself (both the zero-lag triggers and those obtained using timeslides), we obtain a higher significance for GW170818 than our previous result (Venumadhav et al., 2020) (see also (Zackay et al., 2019a)). This also changes the values of the IFAR of the rest of the candidates, as we report in Table 4.A.1.

4.A.2 Search completeness

A binary merger might fail to be detected for a variety of reasons. In our primary search (Venumadhav et al., 2019), a successful detection requires identifying triggers in coincidence at both LIGO detectors, and these coincident triggers must survive a battery of signal quality tests (vetoes). Finally, the candidate has to stand out from the noise background in order to be detected with any significance.

As a diagnosis of the performance of our search pipeline, in Fig. 4.A.1a we show the recovery fraction of injections as a function of their squared injected SNR (defined below), as well as the

estimates are based. In this work we used Virgo data only for parameter estimation of events whenever these are available.

relative frequency of various failure modes. We include in Fig. 4.A.1a only injections with parameters within the target region of our template bank (Roulet et al., 2019), and that happened during times flagged by both the LVC (Abbott et al., 2021) and our pipeline (Venumadhav et al., 2019) as valid for search. Injections labeled 'Missed' (blue) are those that failed to produce a coincident trigger. These include cases when the recovered signal in one of the two LIGO detectors was below our collection threshold ($\rho^2 < 16$), or when a noise transient caused a different template to generate a louder trigger in one of the detectors (some of these second cases could have passed a full 'production' search, since we did not do data-cleaning as part of this injection campaign). 'Vetoed' injections (orange) are those that triggered the signal quality checks, and consequently rejected. We designed these tests to have a false positive rate of a few percent with Gaussian noise for signals having $\rho < 20$ (with an increased rate for much louder events). This is in line with what we observe in Fig. 4.A.1a. Finally, we distinguish injections that we found below or above a moderate inverse false-alarm rate of 10 observing runs in their template bank (of which we searched five). We show these in light and dark gray, respectively; the four curves in Fig. 4.A.1a add up to unity. We derived the fractions using a sliding window in $\rho^2_{\rm inj}$ that averages over 200 contiguous injection samples.

Since we performed independent searches for five different template banks, the same injection might have a different outcome in each search; and likewise at each of the two LIGO detectors. We summarized this information by assigning to each injection the latest stage it got to in the worst detector, and (then) the best bank.

We distinguish between 'injected' and 'recoverable' squared SNR; respectively:

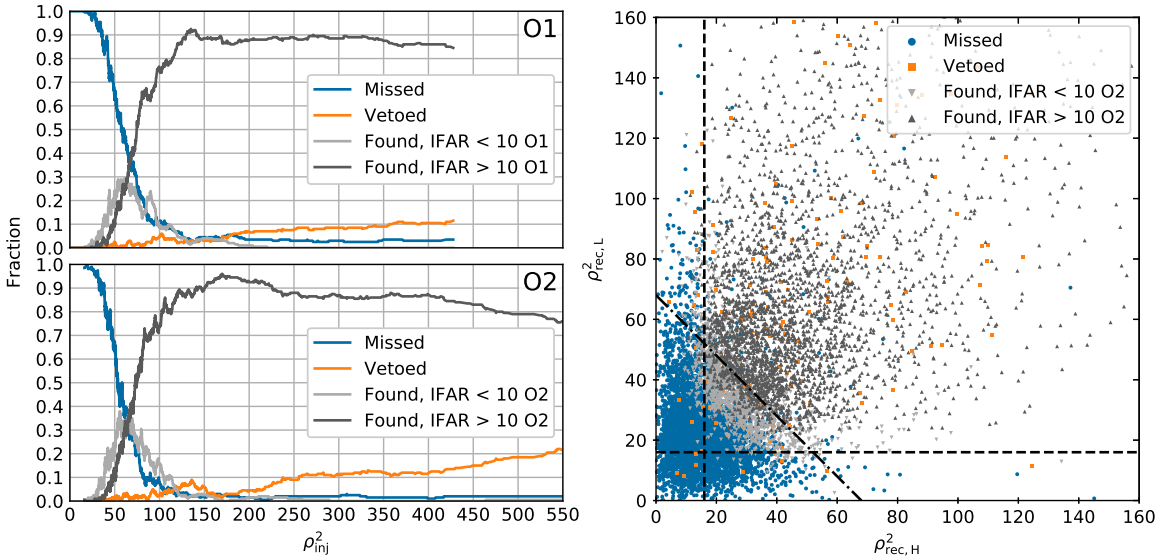
$$\rho^2_{\rm inj} = \sum_{k\in \rm H,L} \langle h \mid h\rangle_k, \tag{4.59}$$

$$\rho^2_{\rm rec} = \sum_{k\in \rm H,L} \langle d \mid h\rangle_k, \tag{4.60}$$

where d and h are the strains of the data and injection respectively. Both $\rho^2_{\rm inj}$ and $\rho^2_{\rm rec}$ depend on the signal parameters as well as the detector sensitivities and orientations at the time of the event. Note that $\rho^2_{\rm inj}$ is independent of the particular noise realization and thus unobservable, unlike $\rho^2_{\rm rec}$. Both are independent of the template bank of the search.

We find that the results for O1 and O2 are comparable. The completeness of the search saturates around 90–95% for signals with $\rho^2_{\rm inj} \gtrsim 150$. At lower SNR values, the dominant failure mode is missed injections, and at higher values it is false rejections (vetoes).

Figure 4.A.1b shows the outcome of the injections made in the O2 run (with the same cuts



(a) Injections in O1 (top) and O2 (bottom).

(b) Injections in O2.

Figure 4.A.1: Probabilities for the possible outcomes of an injection as a function of SNR in our primary search for triggers in Hanford–Livingston coincidence. Only injections with parameters within the target region of our template bank (Roulet et al., 2019), and that happened during times flagged by both the LVC (Abbott et al., 2021) and our pipeline (Venumadhav et al., 2019) as valid for search are included. The four outcomes are exhaustive. *Left panel:* Outcome probabilities as a function of injected squared SNR, for the O1 (top) and O2 (bottom) observing runs. *Right panel:* Outcomes in the O2 observing run in terms of recoverable squared SNRs at the Hanford and Livingston detectors. Dashed lines are drawn at the single-detector collection threshold $\rho^2_{\mathrm{rec\,H,L}} = 16$ and at the approximate Gaussian limit $\rho^2_{\mathrm{rec}} = 68$. Several injections with $\rho^2_{\mathrm{rec}} > 68$ have $\rho^2_{\mathrm{rec,H}} < 16$ due to disparate detector responses (Zackay et al., 2019a).

used for Fig. 4.A.1a) scatter-plotted in the $\rho^2_{\rm rec,L}, \rho^2_{\rm rec,H}$ plane. Recall that by virtue of Eq. (4.25) the injection distribution approximately follows that of astrophysical events in the high SNR limit $\rho^2 > \rho^2_{\rm th} = 60$. The incoherent detection limit of our primary search can be approximated by a Gaussian noise limit $\rho^2_{\rm rec} > 68$ and single-detector collection thresholds $\rho^2_{\rm rec\,H,L} > 16$ (Venumadhav et al., 2020, figure 6a). 60% of the injections with $\rho^2_{\rm rec} > 68$ that are missed from the primary search (or 4.6% of all injections with $\rho^2_{\rm rec} > 68$) have $\rho^2_{\rm rec,H} < 16$, i.e. below our single-detector collection threshold, even though their network SNR is above the Gaussian limit for detection. We performed a targeted search for such signals in (Zackay et al., 2019a), which for computational limitation we do not reproduce here with injections. Such events would stand a second chance of being found in the targeted search, which is not accounted for in Fig. 4.A.1a. Figure 4.A.1b also supports the approximate incoherent detection thresholds for our pipeline used in (Venumadhav et al., 2020, figure 6a).

4.A.3 Sensitive volume–time

Quantifying the sensitive volume–time requires defining a detection threshold, see Eq. (4.18). We will define this threshold in terms of IFAR, that can be compared across different search pipelines. We measure FAR empirically using timeslides, i.e. adding artificial time-shifts between the Hanford and Livingston data streams to generate background triggers, and counting the number of background triggers that have a better detection statistic than a trigger of interest. Our search procedure divides the BBH parameter space into 5 template banks that are explored independently except for the restriction that any trigger is assigned to only one search; further, the O1 and O2 observing runs are analyzed separately. Therefore, the FAR we obtain has units of per bank per $\overline{VT}$ of the relevant observing run.[6] To aid eventual comparisons we express them per all our O1 and O2 BBH searches combined, for which we use (O1 + O2) as notation. For a trigger j:

$$\begin{aligned} \mathrm{FAR}_j &= N_b(\tilde{\rho} > \tilde{\rho}_j)[\mathrm{run}_j\,\mathrm{bank}_j]^{-1} \\ &= N_b(\tilde{\rho} > \tilde{\rho}_j)\frac{(\overline{VT})_{\mathrm{O1+O2}}}{(\overline{VT})_{\mathrm{run}j}} N_{\mathrm{banks}}[\mathrm{O1+O2}]^{-1}, \end{aligned} \tag{4.61}$$

where $\tilde{\rho}$ is the detection statistic of our pipeline and $N_b(\tilde{\rho} > \tilde{\rho}_j)$ is the expected number of background triggers above j estimated from timeslides. We use Eq. (4.61) to aggregate the results from all BBH

[6] This is the reason we report FARs in units of observing runs instead of physical time—the ranking statistic includes a time-dependent volumetric correction factor to account for the significant and systematic changes in the network sensitivity over the run (Venumadhav et al., 2019). If the network sensitivity were constant during the observing run, the units "O1" ≈ 46 days and "O2" ≈ 118 days.

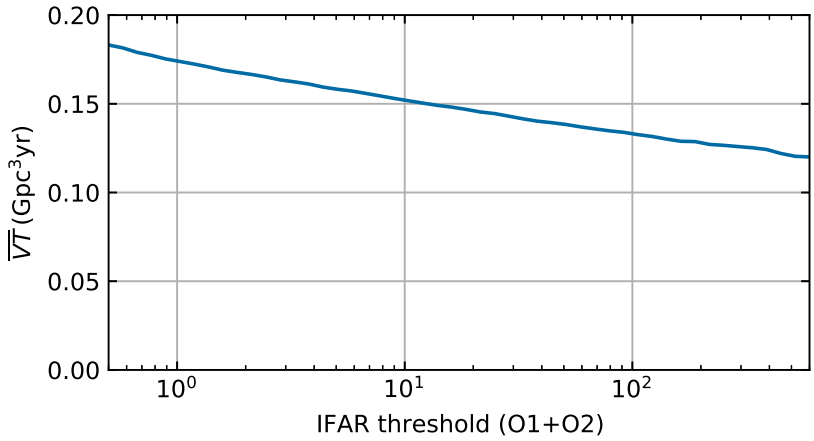


Figure 4.A.2: Sensitive volume–time of our search pipeline on O1 and O2 as a function of IFAR threshold for the default population model defined in Eq. (4.43). The IFAR is referred to the full O1+O2 search and accounts for the fact that multiple template banks were searched per Eq. (4.61).

banks and observing runs. Note that the $\overline{VT}$ estimation in Eq. (4.18) requires a threshold on the FAR, and conversely the FAR in Eq. (4.61) necessitates a computed $\overline{VT}$ ratio between runs. We find a self-consistent solution numerically. Using the default population model defined in Eq. (4.43) and an IFAR threshold of 1(O1 + O2) we obtain $\overline{VT}_{\rm O2}/\overline{VT}_{\rm O1} = 4.13$. This is approximately valid for other population models and thresholds as well since the dependence largely cancels in the ratio. We show the result for $\overline{VT}$ as a function of the IFAR threshold in Fig. 4.A.2, for the default population model of Eq. (4.43).

4.B Computation of the reference $p_{\rm astro}$

As per Eq. (4.7), $p_{{\rm astro},i}$ depends on the ratio of the trigger densities expected from foreground and background locally at the data d_i of each trigger. In principle, d_i consists of the strain time series, or equivalently its Fourier modes. In practice we do not have a reliable model for the background trigger density in terms of these, as we do for the astrophysical events (Eqs. (4.2) and (4.3)). Thus, we instead approximate (for all events except GW170608 and GW170817A, see below):

$$\left.\frac{{\rm d}N_a}{{\rm d}N_b}\right|_{d_i}(\lambda_0) \approx \frac{{\rm d}N_a/{\rm d}\tilde{\rho}^2}{{\rm d}N_b/{\rm d}\tilde{\rho}^2}(\tilde{\rho}_i^2 \mid \mathcal{T} \approx \mathcal{T}_i, \lambda_0), \tag{4.62}$$

where $\tilde{\rho}^2$ is the detection statistic of our pipeline (Venumadhav et al., 2019) and $\mathcal{T}$ identifies the waveform template associated to a trigger. We expect these variables to contain most of the discerning power between astrophysical and background events. The detection statistic is an estimator of the likelihood ratio between the astrophysical and noise hypotheses that incorporates information

from the trigger SNR at Hanford and Livingston, the phase and time differences between detectors and the instantaneous detector sensitivities. It includes only very coarse information about the signal parameters, based on which subbank the triggering template resides in. Since both the astrophysical and background trigger rate densities exhibit dependence on the triggering template, we incorporate the template identity in Eq. (4.62).

We compute Eq. (4.62) for each trigger as follows. We first generate large sets of astrophysical and background triggers. For the former we use the injections described in §4.2.3 and for the latter we use the method of timeslides. We restrict to the triggers that have a similar template to the triggering template $\mathcal{T}_i$: we demand that they are found in the same chirp mass bank (Roulet et al., 2019) and have a match $\langle \mathcal{T} \mid \mathcal{T}_i \rangle$ above some threshold. We choose the highest match threshold that admits at least 100 injection and 100 background triggers.[7] This achieves a compromise between making a density measurement that is local in intrinsic-parameter space and that has an acceptable statistical error. We then do a kernel density estimation of $\mathrm{d}N_a/\mathrm{d}\tilde{\rho}^2$ and $\mathrm{d}N_b/\mathrm{d}\tilde{\rho}^2$, using these triggers with weights (see Eq. (4.25))

$$w_j^{\mathrm{inj}} = \frac{R_0\, Z}{N_{\mathrm{inj}}\, \hat{p}_{\mathrm{det}}(\theta_j)} \tag{4.63}$$

$$w^{\mathrm{bg}} = \frac{1}{N_{\mathrm{timeslides}}} \tag{4.64}$$

for injections and background respectively.

As mentioned in Appendix 4.A, for GW170818 we obtained a $p_{\mathrm{astro}}(\lambda_0) = 0.92$ based on Hanford and Livingston data, higher than reported in (Venumadhav et al., 2020) owing to an improved version of the coherent score, as we described in Sec. 4.A.1. Note the GstLAL pipeline (Sachdev et al., 2019) found $p_{\mathrm{astro}} = 1$ by including Virgo data. As discussed in §4.2.2, following Abbott et al. (2017c) we set $p_{\mathrm{astro}} = 1$ for GW170608.

The event GW170817A was found in a targeted search for signals that are loud in the Livingston detector and faint in the Hanford detector (Zackay et al., 2019a), as opposed to our primary search for signals in Hanford–Livingston coincidence (Venumadhav et al., 2019, 2020). Most of its significance comes from being the loudest Livingston trigger apart from previously confirmed confident signals. As such, the method of timeslides cannot be used to generate empirical background for this event. We compute its $p_{\mathrm{astro}}(\lambda_0)$ following (Zackay et al., 2019a, eqs. (5) and (6)): we define the clean region of parameter space as those templates with chirp mass $\mathcal{M} > 10\,\mathrm{M}_\odot$ for which there were ≤ 5

[7] For bank `BBH 0` (Roulet et al., 2019) we only recovered 24 injections in O2; we just include those in the computation of the p_{astro} of 170412B.

loud Livingston triggers ($\rho_{\rm L}^2 > 60$) in O2 from similar templates (match > 0.9) in times where the Hanford detector was also operating. We obtain the expected number of triggers in this region with $\rho_{\rm L}^2 > 66$ by counting the injections that satisfy all these conditions weighted per Eq. (4.63). We set the expected number of background events to 1.

Recently, Ashton and Thrane (2020) have concluded that the event GW151216 (Zackay et al., 2019b) has $p_{\rm astro} = 0.03$, based on an analysis of background triggers obtained with a different pipeline, as well as foreground triggers generated under an astrophysical prior isotropic in spin directions. A similar method was used by Pratten and Vecchio (2020). Those analyses overlook the fact that different pipelines treat the systematics in the data differently and thus suffer from different backgrounds. As an example, our pipeline applies different data quality checks and signal consistency vetoes. Even within our pipeline, removing or modifying these stages would significantly lower the $p_{\rm astro}$ of near-threshold triggers like GW151216. These tests are not applied in the analysis of (Ashton and Thrane, 2020; Pratten and Vecchio, 2020). They instead characterize the background in terms of its projection onto parameter space when modeled as a GW signal in Gaussian noise. That is a different test that a priori does not have the same discerning power between signals and noise. Ultimately, the choices that maximize the pipeline sensitivity (see Appendix 4.A) should be pursued. As a result of these considerations, $p_{\rm astro}$ is inherently a pipeline-dependent quantity. In addition, $p_{\rm astro}$ depends on the astrophysical population model. As shown in Table 4.3.1, for GW151216 (as well as other near-threshold events with nonzero spin) $p_{\rm astro}$ is particularly sensitive to the spin distribution. Ashton and Thrane (2020); Pratten and Vecchio (2020) used an isotropic spin model, under which the astrophysical interpretation is indeed strongly disfavored. We emphasize however that for different spin models GW151216 has a sizable $p_{\rm astro}$. We should consistently account for this dependence when interpreting GW151216 and assessing its implications for the astrophysical BBH population.

4.C Differences with Galaudage et al. (2020)

Galaudage et al. (2020) have presented a framework similar but inequivalent to the one presented in Section 4.2. In this appendix we compare both treatments and identify the differences between them. Our point of comparison is the model likelihood, our Eq. (4.8) or their equation (36), which are not compatible.

The correspondence between their notation and ours is as follows. Their astrophysical hypothesis

prior is

$$\xi \equiv \frac{N_a}{N_a + N_b} \tag{4.65}$$

in our notation. Their signal likelihood is

$$\mathcal{L}(d \mid \Lambda, \mathrm{det}) = \frac{\mathcal{V}_{\mathrm{tot}}}{\mathcal{V}(\Lambda)} \mathcal{L}(d \mid \Lambda) \equiv \frac{1}{N_a} \frac{\mathrm{d}N_a}{\mathrm{d}d} = \frac{1}{\overline{VT}(\lambda')} \int \mathrm{d}\theta f(\theta \mid \lambda') P(d \mid \theta), \tag{4.66}$$

where their $\Lambda \equiv \lambda'$, $\mathcal{V}(\Lambda) \equiv \overline{VT}(\lambda')$ and we have used Eqs. (4.2), (4.10) and (4.14). Their noise likelihood is

$$\begin{aligned} \mathcal{L}(d \mid \varnothing, \mathrm{det}) &\equiv \frac{1}{N_b} \frac{\mathrm{d}N_b}{\mathrm{d}d} \\ &= \frac{N_a}{N_b} \cdot \frac{\mathrm{d}N_b}{\mathrm{d}N_a} \cdot \frac{1}{N_a} \frac{\mathrm{d}N_a}{\mathrm{d}d} \\ &= \frac{\xi}{1-\xi} \cdot \frac{1 - p_{\mathrm{astro}}}{p_{\mathrm{astro}}} \cdot \mathcal{L}(d \mid \Lambda, \mathrm{det}), \end{aligned} \tag{4.67}$$

using Eqs. (4.7), (4.65) and (4.66). Equation (4.67) is (Galaudage et al., 2020, eq. (7)) generalized to account for selection effects.

In their equation (29), Galaudage et al. (2020) use a prescription for the normalization of the background term in the likelihood, $p_{\varnothing} = \mathcal{V}(\Lambda_0)/\mathcal{V}_{\mathrm{tot}}$, which holds for Λ on the vicinity of Λ_0. Instead, the exact expression that satisfies Eq. (4.67) is

$$p_{\varnothing}(\Lambda) = \frac{\mathcal{V}(\Lambda)}{\mathcal{V}_{\mathrm{tot}}}. \tag{4.68}$$

In (Galaudage et al., 2020, eq. (35)) it is stated that $(p_{\mathrm{astro}}^{-1} - 1)\mathcal{L}(d \mid \Lambda)$ is independent of Λ, so they equate this term to its value for a fiducial Λ_0. However, p_{astro} also depends on the rate of mergers, so the fiducial model should include the rate (encoded in ξ). Additionally, their equation (35) relies on equation (7) which does not include selection effects. The correct form of (Galaudage et al., 2020, eq. (35)) is thus

$$\frac{\xi}{1-\xi} \frac{1 - p_{\mathrm{astro}}(\xi, \Lambda)}{p_{\mathrm{astro}}(\xi, \Lambda)} \frac{\mathcal{L}(d \mid \Lambda)}{\mathcal{V}(\Lambda)} = \frac{\xi_0}{1-\xi_0} \frac{1 - p_{\mathrm{astro}}(\xi_0, \Lambda_0)}{p_{\mathrm{astro}}(\xi_0, \Lambda_0)} \frac{\mathcal{L}(d \mid \Lambda_0)}{\mathcal{V}(\Lambda_0)}, \tag{4.69}$$

where ξ_0, Λ_0 are the values used to compute the fiducial p_{astro}.

Lastly, in (Galaudage et al., 2020, eq. (34)) the expression $\xi = N/n$ is used, where $N \equiv N_{\mathrm{trig}}$ is the number of triggers and n the number of data segments analyzed. But ξ encodes a (model dependent) prior expectation of the ratio of counts of astrophysical and noise events, and not the

actual outcome of the experiment. The correct expression is instead

$$\xi = \frac{R\mathcal{V}(\Lambda)}{R\mathcal{V}(\Lambda) + R_g T_{\rm obs}}. \tag{4.70}$$

Once the changes in Eqs. (4.68), (4.69) and (4.70) have been applied, their equation (36) becomes

$$\begin{aligned} &P(\{d_i\}, N \mid \Lambda) \\ &\quad = \frac{e^{-(R\mathcal{V}(\Lambda)+R_g T_{\rm obs})}(R\mathcal{V}(\Lambda))^N}{N!} \prod_i \frac{\mathcal{V}_{\rm tot}}{\mathcal{V}(\Lambda)} \frac{\mathcal{L}(d_i \mid \Lambda)}{p_{{\rm astro},i}(R,\Lambda)} \\ &\quad = \frac{e^{-(R\mathcal{V}(\Lambda)+R_g T_{\rm obs})}(R_0\mathcal{V}_{\rm tot})^N}{N!} \prod_i \left(\frac{1 - p_{{\rm astro},i}(R_0,\Lambda_0)}{p_{{\rm astro},i}(R_0,\Lambda_0)} \frac{R_g}{R_{g,0}} + \frac{R}{R_0} \frac{\mathcal{L}(d_i \mid \Lambda)}{\mathcal{L}(d_i \mid \Lambda_0)} \right) \mathcal{L}(d_i \mid \Lambda_0), \end{aligned} \tag{4.71}$$

which is compatible with Eq. (4.8). The rate R does not factor out, so we cannot marginalize it analytically in Eq. (4.71) without expanding the binomial first.

As a consequence of Eq. (4.69),

$$p_{\rm astro}(\xi,\Lambda) = \left(\frac{\xi_0}{1-\xi_0} \frac{\mathcal{L}(d \mid \Lambda_0)}{\mathcal{V}(\Lambda_0)} \frac{1-\xi}{\xi} \frac{\mathcal{V}(\Lambda)}{\mathcal{L}(d \mid \Lambda)} + 1 \right)^{-1}, \tag{4.72}$$

instead of (Galaudage et al., 2020, eq. (42)). This difference might explain why Table 4.3.1 does not reproduce the results of (Galaudage et al., 2020, table II). We emphasize that the fiducial $p_{\rm astro}$ have to correspond to R_0, Λ_0 and reported values must be interpreted with this in mind. In Table 4.3.1 we have computed the $p_{\rm astro}$ of the top triggers our pipeline found under a specific astrophysical model λ_0 to facilitate this task.

4.D Robustness of the reweighting procedure

Our framework involves Monte Carlo computation of integrals by reweighting samples: the numerator and denominator of Eq. (4.17) use source parameter estimation samples, and Eq. (4.18)) uses injection samples. These methods are subject to stochastic errors, especially if the target and proposal distributions are mismatched. We have chosen the proposal distributions with this consideration in mind, in this appendix we show that our procedure indeed achieved sufficient robustness.

We use the bootstrap method to estimate the impact that stochastic error in these integrations has on the population inference. For each population model, we do 100 bootstrap repetitions of the inference of population parameters λ, each time employing a set of samples taken randomly with

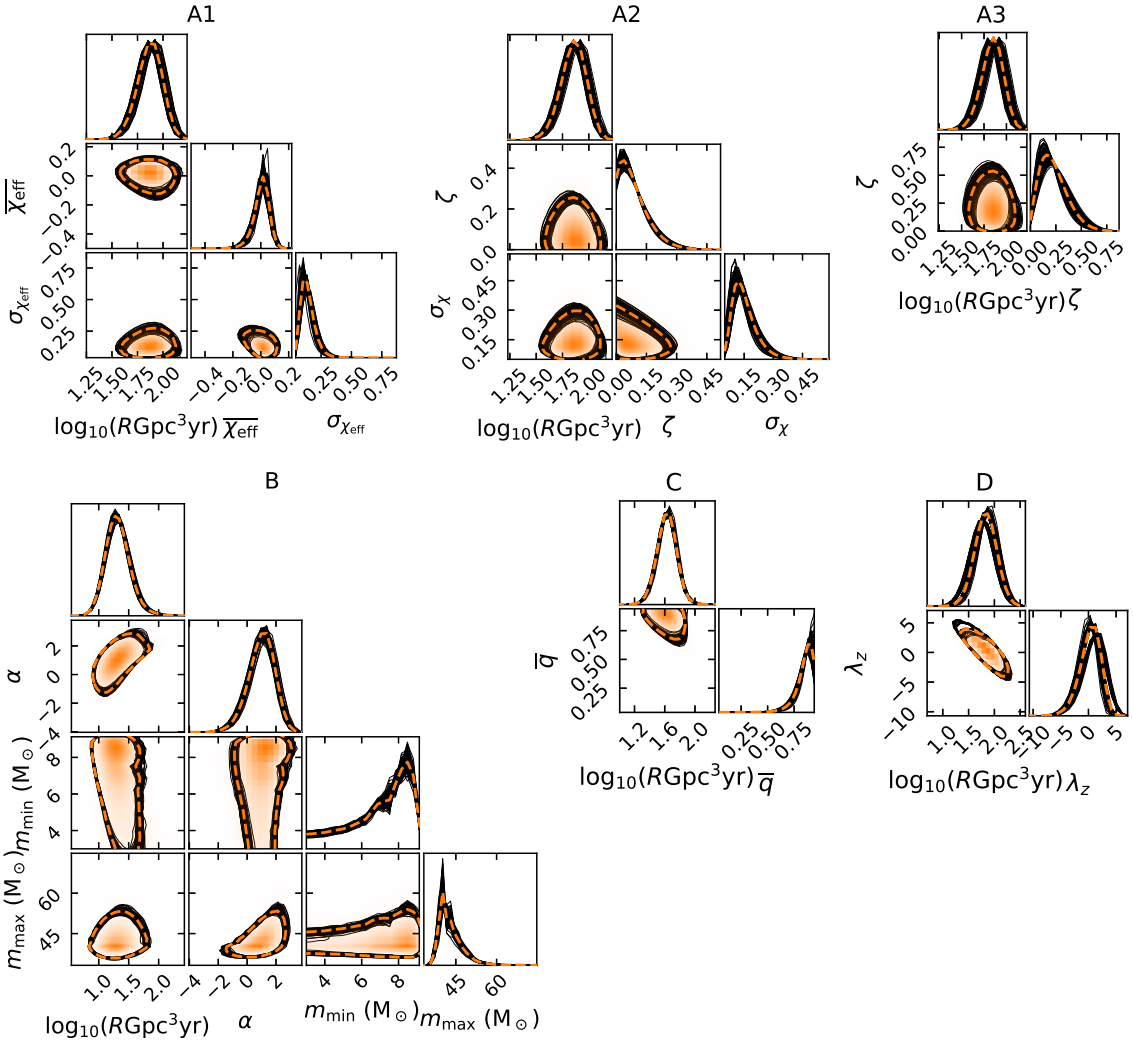


Figure 4.D.1: 100 bootstrap realizations for the population Models A1–D explored in this work (thin black), along with our results from §4.3.1–4.3.4 (dashed orange). These quantify the impact of stochastic errors associated with the various Monte Carlo integrations involved, which remains small at the current uncertainty level. Contours enclose 90% of the distribution.

replacement from the corresponding original sets for all Monte Carlo integrations involved—i.e., from the source parameter estimation samples of each event for the computation of $\mathcal{W}_i(\lambda' \mid \lambda'_0)$ in Eq. (4.17), and from the set of injections for $\overline{\mathcal{VT}}(\lambda')$ in Eq. (4.18). The collection of these inference realizations may be interpreted as the ensemble of solutions compatible with our stochastic sampling errors. Figure 4.D.1 shows the results of this exercise for Models A1–D and the combined GWTC-1 and IAS catalog. Thin black lines show the 100 bootstrap realizations and dashed orange lines show our original results. These stochastic errors are smaller than the uncertainty levels, providing evidence that the proposal distribution choices and the number of injections and parameter estimation samples were adequate to achieve a robust reweighting.

Chapter 5

Effective Spins and Masses from O1–O3a

This chapter has been adapted from Roulet et al. (2021).

The distribution of effective spin $\chi_{\rm eff}$, a parameter that encodes the degree of spin–orbit alignment in a binary system, has been widely regarded as a robust discriminator between the isolated and dynamical formation pathways for merging binary black holes. Until the recent release of the GWTC-2 catalog, such tests have yielded inconclusive results due to the small number of events with measurable nonzero spins. In this work, we study the $\chi_{\rm eff}$ distribution of the binary black holes detected in the LIGO–Virgo O1–O3a observing runs. Our focus is on the degree to which the $\chi_{\rm eff}$ distribution is symmetric about $\chi_{\rm eff} = 0$ and whether the data provides support for a population of negative-$\chi_{\rm eff}$ systems. We find that the $\chi_{\rm eff}$ distribution is asymmetric at 95% credibility, with an excess of aligned-spin binary systems ($\chi_{\rm eff} > 0$) over anti-aligned ones. Moreover, we find that there is no evidence for negative-$\chi_{\rm eff}$ systems in the current population of binary black holes. Thus, based solely on the $\chi_{\rm eff}$ distribution, dynamical formation is disfavored as being responsible for the entirety of the observed merging binary black holes, while isolated formation remains viable. We also study the mass distribution of the current binary black hole population, confirming that a single truncated power law distribution in the primary source-frame mass, $m_{1\rm s}$, fails to describe the observations. Instead, we find that the preferred models have a steep feature at $m_{1\rm s} \sim 40\,{\rm M}_\odot$ consistent with a step and an extended, shallow tail to high masses.

5.1 Introduction

The growing number of gravitational wave sources observed by the LIGO and Virgo detectors is leading to an improved picture of the astrophysical population of binary mergers. The recent release of the second Gravitational-Wave Transient Catalog, GWTC-2 (Abbott et al., 2020e), by the LIGO–Virgo Collaboration (LVC) has roughly tripled the sample size of observed binary black hole mergers (Abbott et al., 2019a; Nitz et al., 2019a; Venumadhav et al., 2019; Zackay et al., 2019b; Venumadhav et al., 2020; Zackay et al., 2019a; Nitz et al., 2020) and is starting to offer hints about the astrophysical origin of these binary systems (Abbott et al., 2020f; Wong et al., 2021; Zevin et al., 2021; Bouffanais et al., 2021).

Indeed, the distribution of binary black hole parameters (e.g. masses, spins, redshift) is an observable that allows us to test models of formation pathways for these systems. Proposed scenarios include dynamical assembly and hardening of binary black holes in dense stellar environments, such as globular clusters (Zwart and McMillan, 1999; O'Leary et al., 2006; Sadowski et al., 2008; Downing et al., 2010, 2011; Samsing et al., 2014; Rodriguez et al., 2015, 2016a; Askar et al., 2016), nuclear star clusters (Antonini and Rasio, 2016; Petrovich and Antonini, 2017), and young stellar clusters (Ziosi et al., 2014; Mapelli, 2016; Banerjee, 2017; Chatterjee et al., 2017); isolated evolution of a binary star in the galactic field, which undergoes either a common envelope phase (Nelemans et al., 2001; Belczynski et al., 2002; Voss and Tauris, 2003; Belczynski et al., 2007, 2008; Dominik et al., 2013; Belczynski et al., 2014; Mennekens and Vanbeveren, 2014; Spera et al., 2015; Eldridge and Stanway, 2016; Stevenson et al., 2017b; Mapelli et al., 2017; Giacobbo et al., 2017; Mapelli and Giacobbo, 2018; Kruckow et al., 2018; Giacobbo and Mapelli, 2018) or a chemically homogeneous evolution (Marchant et al., 2016; de Mink and Mandel, 2016; Mandel and de Mink, 2016); and binary mergers prompted by interactions with a supermassive black hole (Antonini and Perets, 2012), gas and stars in the accretion disk of an active galactic nucleus (McKernan et al., 2012; Stone et al., 2016; Bartos et al., 2017), or additional companions in higher-multiplicity systems (Antonini et al., 2014; Kimpson et al., 2016; Antonini et al., 2017; Liu and Lai, 2018; Hamers et al., 2015).

Since the individual components of the dimensionless spin vectors $\boldsymbol{\chi_1}$ and $\boldsymbol{\chi_2}$ are hard to measure (Vitale et al., 2014; Pürrer et al., 2016; Vitale et al., 2017a) and their directions generally evolve with time due to precession (Apostolatos et al., 1994; Kidder, 1995), a well-known effective aligned-spin parameter was introduced (Racine, 2008; Ajith et al., 2011; Santamaria et al., 2010)

$$\chi_{\rm eff} := \frac{\boldsymbol{\chi_1} + q\boldsymbol{\chi_2}}{1+q} \cdot \hat{\boldsymbol{L}}, \tag{5.1}$$

where $\hat{\boldsymbol{L}}$ is the unit vector along the Newtonian orbital angular momentum of the binary, $q = m_2/m_1 \leq 1$ is the mass ratio. The effective spin is motivated by the fact that it can be measured relatively precisely in the data, and is approximately conserved throughout the binary coalescence after orbit averaging (Racine, 2008). No less important, two of the main broad classes of binary black hole formation channels make predictions about qualitative features of the effective spin distribution that are robust to model uncertainties. Dynamical formation channels in general predict that the spins and orbit should be isotropically distributed and uncorrelated with each other. In particular, this implies that for these systems the $\chi_{\rm eff}$ distribution is symmetric about 0. Isolated formation channels instead predict correlations in the spins and orbit directions due to mass transfer episodes or tidal interactions between the component stars. As a result, the isolated scenario predicts a distribution of $\chi_{\rm eff}$ with little support at negative values. Within this channel, a small fraction of mergers with negative $\chi_{\rm eff}$ could still possibly be explained by anisotropic supernova explosions at the black holes formation, which impart a natal kick that can change the plane of the orbit and thus the value of $\chi_{\rm eff}$ (Rodriguez et al., 2016b; Gerosa et al., 2018). However, if these kicks were strong enough to invert the direction of the orbit in a sizeable fraction of the cases, they would also unbind the binaries so frequently that the observed rates would be hard to explain (Belczynski and Bulik, 1999; Callister et al., 2020).

In this work we will study in detail the degree to which these two qualitative features of the $\chi_{\rm eff}$ distribution, namely its symmetry about 0 and support at negative values, hold for the observed sample. Both features become hard to test if black hole spins are small, which is predicted from stellar evolution models (Fuller and Ma, 2019; Bavera et al., 2020) and is also the case of most observations. Indeed, until the recent release of the GWTC-2 catalog these simple but general tests were mostly inconclusive due to the small number of events with measurable nonzero $\chi_{\rm eff}$ (Farr et al., 2017, 2018; Roulet and Zaldarriaga, 2019; Roulet et al., 2020). Including events from the O3a observing run, Abbott et al. (2020f) first reported evidence for both features in the population: they found the $\chi_{\rm eff}$ distribution to have a positive mean and support at negative values. Together, these observations suggest that neither dynamical nor isolated formation channels alone can explain the entirety of the detections. Combining this information with the observed mass distribution, Zevin et al. (2021); Bouffanais et al. (2021) reached a similar conclusion, and applied a further layer of interpretation to constrain uncertain parameters of physical models of binary black hole formation. Here, we instead constrain a phenomenological description of the binary black hole population, more akin to the analysis of Abbott et al. (2020f).

The mass distribution is another observable that can inform binary black hole formation chan-

nels, as well as physical processes of stellar evolution. Of special interest is the high-mass end of the mass distribution observable by LIGO–Virgo, $m \gtrsim 40\,\mathrm{M}_\odot$. Due to the (pulsational) pair instability supernova process, black holes with mass between $\sim 45\,\mathrm{M}_\odot$ and $135\,\mathrm{M}_\odot$ are not expected to form from stellar collapse ("upper mass gap") (Fowler and Hoyle, 1964; Barkat et al., 1967; Bond et al., 1984; Heger et al., 2003; Farmer et al., 2019). A natural way to produce black holes in this mass range is through mergers of lighter black holes. In dense environments these so-called "second-generation" black holes can become paired and merge again, emitting an observable gravitational wave signal. This process is contingent on retention of the remnant black hole, so its efficiency depends on the interplay between the merger kick (a recoil of the remnant black hole due to asymmetric gravitational wave emission at merger) and the local escape velocity. The magnitude of the kicks is sensitive to the spins of the merging black holes, smaller spins usually yielding smaller kicks. In turn, different types of dense environments have different escape velocities, typical numbers being 10–$10^2\,\mathrm{km\,s^{-1}}$ for globular clusters and up to $\sim 10^3\,\mathrm{km\,s^{-1}}$ for nuclear clusters. Second generation mergers do not happen for binaries formed in isolation. Some alternative pathways to produce black holes in this mass range may involve accretion of gas (Safarzadeh et al., 2020) or extreme values of the $^{12}\mathrm{C}(\alpha,\gamma)^{16}\mathrm{O}$ nuclear cross section, which can shift the location of the mass gap (Farmer et al., 2019), see (Gerosa and Fishbach, 2021) and references therein for a recent review. On the observational side, current interferometers are particularly sensitive to mergers in this high-mass region of parameter space, which makes it a promising discriminator (Fishbach and Holz, 2017). Indeed, some events were observed to have significant support for one or both component black holes in this mass range (e.g. GW190521, GW190602_175927, GW190706_222641, GW190519_153544, GW190929_012149 (Abbott et al., 2020e), GW170817A (Zackay et al., 2019a)). While analyses prior to O3a found evidence for a cut-off in the mass distribution at $\sim 40\,\mathrm{M}_\odot$ (Fishbach and Holz, 2017; Wysocki et al., 2019; Roulet and Zaldarriaga, 2019; Roulet et al., 2020), this picture changed with the inclusion of O3a and models with more structure, including a tail at high mass, became favored (Abbott et al., 2020f). Here, we will also explore parametric models of the primary mass distribution in order to validate these results.

Our main findings are:

1. The χ_{eff} distribution is inconsistent with being symmetric about zero at the 95% credible level, with aligned-spin binary systems ($\chi_{\mathrm{eff}} > 0$) predominating over those with anti-aligned spins ($\chi_{\mathrm{eff}} < 0$). This result provides some evidence against the formation scenario in which the entire population of binary black holes has isotropically-distributed spins, as predicted if all

merging binary black holes are formed dynamically in dense stellar environments;

2. We find no evidence for negative $\chi_{\rm eff}$ in the population, in contrast to Abbott et al. (2020f). We are able to reproduce their results, but find that the parametrized model they used in order to reach this conclusion is disfavored by the data and that the inferred presence of negative spins is contingent on this parametrization;

3. We find that the primary-mass distribution steepens at $\sim 40\,\mathrm{M}_\odot$ and then flattens, with an extended tail to high masses whose detailed shape is hard to constrain with current data.

This paper is organized as follows: in Section 5.2, we describe the data investigated in this work, our sample selection criteria, and the parameter estimation method used to infer the source parameters of the binary black holes. In Section 5.3, we conduct a model-free exploration of the data, with a special focus on the empirical distribution of $\chi_{\rm eff}$. In Section 5.4, we describe our statistical methods for model selection and apply them to several parametrized models for the distributions of the effective spin and primary mass. We conclude in Section 5.5. We provide details of the sample of events that we use in Appendix 5.A.

5.2 Data

The data explored in this work consists of the binary black hole events reported in the LVC GWTC-1 (Abbott et al., 2019a) and GWTC-2 (Abbott et al., 2020e) catalogs, and those identified in the independent IAS O1–O2 catalog (Venumadhav et al., 2019; Zackay et al., 2019b; Venumadhav et al., 2020; Zackay et al., 2019a). Some of the events reported in the IAS O1–O2 catalog have been independently confirmed by Nitz et al. (2019a, 2020). Following the main analysis conducted by the LVC in their population study (Abbott et al., 2020f), we exclude GW190814 (Abbott et al., 2020d) in this work as it is an outlier with respect to the rest of the observed population, and for ease of comparison between our results and the LVC's results (see Section 5.4). We do not include the recent 3-OGC catalog (Nitz et al., 2021), which was published as this work was being completed. A summary of the events used in this work is provided in Appendix 5.A.

Depending on how easily our models for astrophysical signals and detector noise can account for the properties of a given trigger, some detections are more statistically significant than the others. Roulet et al. (2020) provided a framework to take this into account when using triggers of arbitrary significance. However, in order to simplify the interpretation of the results shown in Section 5.3, we find it convenient to define a "gold sample" of events that are confidently astrophysical in origin. For

a similar reason, we also exclude from the gold sample those events that happened when a detector exhibited non-Gaussian noise transients, which makes estimation of their parameters and significance more challenging. We include an event in the gold sample if (i) it was identified by at least two search pipelines with a false-alarm rate FAR $< 0.1\,\mathrm{yr}^{-1}$; and (ii) none of the detectors exhibited non-Gaussian transient noise in its vicinity (see Appendix 5.A for details). These criteria are neither explicitly dependent on nor expected to correlate significantly with the binary black hole intrinsic parameters; as such, our gold sample constitutes an unbiased representation of the distribution for the intrinsic parameters of detectable mergers. Indeed, as we shall see in Section 5.4, our conclusions are not strongly affected by this choice of sample. Out of the total 55 events considered in this work, 30 are in the gold sample (see Appendix 5.A).

We infer the source parameters of each binary system with the `IMRPhenomXPHM` model, which describes the gravitational waves emitted by a quasi-circular binary black hole (Pratten et al., 2020). This model accounts for spin–orbit precession and the $(\ell, |m|) = \{(2,2),(2,1),(3,3),(3,2),(4,4)\}$ harmonics of the gravitational radiation. We use the relative binning algorithm to evaluate the likelihood (Zackay et al., 2018), and `PyMultiNest` (Buchner et al., 2014) to sample the posterior distribution. For the events identified near non-Gaussian transient noise (summarized in Ref. (Abbott et al., 2020e, table V)), we do not make special mitigation efforts, though we verify that we obtain parameter estimation results that are similar to those reported by Abbott et al. (2020e), who applied glitch subtraction algorithms before performing parameter estimations (Cornish and Littenberg, 2015; Littenberg et al., 2016; Cornish et al., 2021).

For each event, we sample the posterior distribution using a prior that is uniform in detector-frame component masses, χ_{eff} and luminosity volume. For the remaining spin components, we adopt a uniform prior for the poorly-measured variable $\chi_{\mathrm{diff}} := (q\boldsymbol{\chi_1} - \boldsymbol{\chi_2}) \cdot \hat{\boldsymbol{L}}/(1+q)$, conditioned on χ_{eff} and enforcing the Kerr limit on the individual spin magnitudes, $|\boldsymbol{\chi_1}| \leq 1$ and $|\boldsymbol{\chi_2}| \leq 1$. χ_{eff} and χ_{diff} together determine the two spin components that are aligned with the orbital angular momentum, χ_{1z} and χ_{2z}. We then take the prior of the in-plane spin components of the black holes, χ_{ix} and χ_{iy} with $i = 1, 2$, to be uniformly distributed in the disk $\chi_{ix}^2 + \chi_{iy}^2 \leq 1 - \chi_{iz}^2$.

Our parameter estimation results are broadly consistent with LVC's after accounting for the differences in spin priors, with two notable exceptions. Firstly, we find that the posterior distribution for GW151226 (Abbott et al., 2016b) significantly changes towards more unequal mass ratio, larger positive χ_{eff} and more misaligned primary spin when higher harmonics and precession are included in the parameter estimation (Chia et al., 2021). Another remarkable event is GW190521, which was reported to have component source-frame masses $m_{1\mathrm{s}} = 85^{+21}_{-14}\,\mathrm{M}_\odot$, $m_{2\mathrm{s}} = 66^{+17}_{-18}\,\mathrm{M}_\odot$ by the

LVC (Abbott et al., 2020c). Using a different prior for the masses and distance, and allowing for a broader parameter range, Nitz and Capano (2021) found a qualitatively different trimodal solution, with roughly similar total mass and peaks at $q \sim 1/2, 1/5$, and $1/12$. Instead, we find a bimodal solution which is approximately consistent with the first two of these modes (Olsen et al.), similar to that reported in (Estellés et al., 2021).

5.3 Model-free exploration

In this section, we carry out a model-free exploration of the data. Our emphasis is on the symmetry, or lack thereof, between positive and negative values of $\chi_{\rm eff}$ in the observed $\chi_{\rm eff}$ distribution. We also investigate if the data requires a distribution with support at negative values of $\chi_{\rm eff}$. To ease the interpretation of the plots shown in this section, we shall restrict ourselves to the events identified in the gold sample (see Section 5.2). We defer a model-dependent analysis of the data to Section 5.4.

5.3.1 Support for nonzero $\chi_{\rm eff}$

We first test the simplest hypothesis that all binary black holes have $\chi_{\rm eff} = 0$, with any apparent deviation away from zero arising due to measurement uncertainty. This test is motivated by the fact that, while the $\chi_{\rm eff}$ measurements of some of the events have most of their support at $\chi_{\rm eff} < 0$ (GW170121, GW150914, GW170818, GW190421_213856, GW170104, GW151012, GW190915_235702, GW170727, GW190521, GW190408_181802), none of them confidently excludes $\chi_{\rm eff} = 0$. In the left panel of Fig. 5.3.1, we explore whether the observed scatter in the $\chi_{\rm eff}$ distribution is consistent with noisy measurements of a $\chi_{\rm eff} = 0$ population. We plot the empirical distribution of the quantity $\langle\chi_{\rm eff}\rangle/\sigma$, i.e. the mean $\chi_{\rm eff}$ of each of the event's posterior samples divided by their standard deviation, and compare it with the cumulative of a standard Gaussian distribution with zero mean and amplitude N_0, where N_0 is the number of events in this distribution. Provided that the likelihood is approximately Gaussian as a function of $\chi_{\rm eff}$, these distributions should match if the true $\chi_{\rm eff}$ were 0. In particular, with the current number of observed events, we would not expect to find events that are more than 2σ away from $\chi_{\rm eff} = 0$. In the left panel of Fig. 5.3.1, we observe that although $N_0 \approx 20$ out of the 30 events in the gold sample are consistent with noisy measurements of a $\chi_{\rm eff} = 0$ distribution, there is an excess of about 10 events with $\chi_{\rm eff} > 0$ that cannot be explained by measurement uncertainty. On the other hand, no such tail seems to be needed in the $\chi_{\rm eff} < 0$ interval.

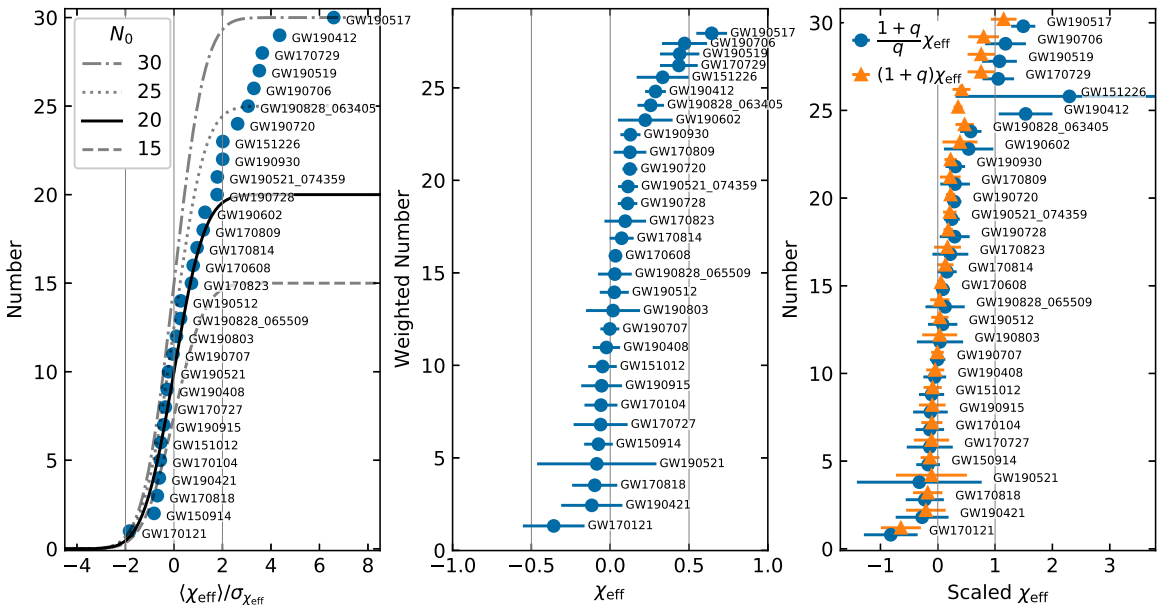


Figure 5.3.1: Empirical spin distributions of the events in the gold sample (see Appendix 5.A). For each event, parameter estimation samples were obtained using a uniform prior in $\chi_{\rm eff}$. To avoid clutter, event names were abbreviated when this did not cause ambiguity. *Left panel:* Mean effective spin scaled by the standard deviation for each event's posterior. We see that about $N_0 \approx 20$ events in the gold sample are consistent with noisy measurements of a $\chi_{\rm eff} = 0$ subpopulation, but a tail in the positive $\chi_{\rm eff}$ end of the distribution is clearly needed in order to accommodate the remaining ≈ 10 events. Conversely, no such tail seems to be needed at the negative end. *Middle panel:* $\chi_{\rm eff}$ distribution, where markers and error bars indicate mean and standard deviation. In the cumulative, each event is weighted by the ratio of the event's sensitive volume to a similar event with zero spins in order to cancel spin selection effects. We see that there are several events with small but well-measured $\chi_{\rm eff} > 0$ for which spin selection effects are not important. *Right panel:* ratio of observed $\chi_{\rm eff}$ to its characteristic value if strong tides were acted on the secondary (blue circles) or primary (orange triangles) black hole progenitor. A number of events are inconsistent with any of these variables being 0 or 1, thereby excluding the strong-tide model as the only mechanism generating black hole spins.

5.3.2 Symmetry of the $\chi_{\rm eff}$ distribution

The observed excess of $\chi_{\rm eff} > 0$ events relative to those with $\chi_{\rm eff} < 0$ in the left panel of Fig. 5.3.1 does not immediately imply that the astrophysical $\chi_{\rm eff}$ distribution is asymmetric about $\chi_{\rm eff} = 0$: an important caveat is that, other parameters being equal, mergers with large and positive values of $\chi_{\rm eff}$ are louder due to the so-called "orbital hangup" effect (Campanelli et al., 2006). This effect leads to a selection bias favoring more observations of $\chi_{\rm eff} > 0$ events, even if the astrophysical $\chi_{\rm eff}$ distribution is symmetric (Ng et al., 2018; Roulet and Zaldarriaga, 2019). The observed excess of $\chi_{\rm eff} > 0$ events thus requires careful interpretation.[1] In the middle panel of Fig. 5.3.1, we plot the

[1]Strictly speaking, this observational bias is also present in the left panel of Fig. 5.3.1 through the parameter estimation prior: in order to match the observed amplitude of a signal, higher-$\chi_{\rm eff}$ solutions are located farther out in distance and thus have more phase space volume available. In other words, a flat prior for the astrophysical $\chi_{\rm eff}$ distribution is implicitly skewed towards positive $\chi_{\rm eff}$ values when conditioned on the strain amplitude measured at

empirical distribution of $\chi_{\rm eff}$ correcting for this observational bias. The bias is computed as follows: for each event, we compute the weight factor

$$w = \langle V_{\rm no\,spin}/V \rangle, \tag{5.2}$$

which is inversely proportional to the sensitive volume V to the corresponding event. Here, we approximate V of a source that has (detector-frame) intrinsic parameters $\theta_{\rm int} = \{m_1, m_2, \chi_{1z}, \chi_{2z}\}$ as

$$V(\theta_{\rm int}) \propto \rho_0^3(\theta_{\rm int}), \tag{5.3}$$

with ρ_0 the single-detector signal-to-noise ratio (SNR) of an overhead, face-on source at a fiducial luminosity distance with a fiducial sensitivity. $V_{\rm no\,spin}$ is defined similarly but with $\chi_{1z} = \chi_{2z} = 0$. For simplicity, we set the in-plane spin components to zero and neglect cosmological evolution throughout this computation. We then average the ratio of these two volumes over the posterior distribution of each event in order to obtain the weight w. In the middle panel of Fig. 5.3.1, we see that many of the events that deviate most significantly away from $\chi_{\rm eff} = 0$ have small values of $\chi_{\rm eff}$ and hence a small impact in the sensitive volume. In particular, GW190728_064510, GW190521_074359, GW190720_000836, GW190930- _133541, GW190828_063405 and GW190412 are $\gtrsim 2\sigma$ away from $\chi_{\rm eff} = 0$ and have relatively small values of $\chi_{\rm eff} \sim 0.1$–0.25. The vertical spacing between events in this plot is given by the volume weight w of the event: for the first four of these events w is approximately 0.95, and for the last two approximately 0.75. Since these are small volume corrections, there is no compelling reason as to why the same number of corresponding events on the negative side of $\chi_{\rm eff}$ should not be seen, for an astrophysical $\chi_{\rm eff}$ distribution that is symmetric about zero. There are also events that are more than 2σ away from zero $\chi_{\rm eff}$ but with relatively large values of $\chi_{\rm eff} \sim 0.5$ (GW170729, GW190519_153544, GW190706_222641 and GW190517_055101). The selection effect for these events is more appreciable, $w \sim 0.5$, so it would be easier to miss similar events with the opposite sign of $\chi_{\rm eff}$. Altogether, the left and middle panels of Fig. 5.3.1 hint that the empirical effective spin distribution is consistent with a distribution with no support for negative spins, but not so much with one symmetric about $\chi_{\rm eff} = 0$.

the detector. As a result, noisy measurements of a $\chi_{\rm eff} = 0$ distribution would be slightly biased towards $\chi_{\rm eff} > 0$. However, within that scenario this effect is small, as the measured $\chi_{\rm eff}$ would be small.

5.3.3 Testing tidal models

Finally, we explore whether the observed events with positive $\chi_{\rm eff}$ can be explained by a simple model of tides acting on the progenitor of one of the component black holes. The simplest and most extreme model for tides assumes that tides sourced by the companion are either very efficient at spinning up the progenitor star or negligible depending on the orbital separation after a common envelope phase, because tidal torques are very sensitive to the orbital separation. Then, a fraction of the component black holes would come from tidally-torqued progenitors and would have a large, aligned spin $\chi_z \approx 1$ (Kushnir et al., 2016; Zaldarriaga et al., 2017). If, barring tides, natal black hole spins were small (Fuller and Ma, 2019; Bavera et al., 2020), the $\chi_{\rm eff}$ distribution would have peaks at $\chi_{\rm eff} = 0, q/(1+q), 1/(1+q), 1$ when the tides were inefficient, torqued the progenitor of the secondary black hole, torqued the progenitor of the primary, or torqued both, respectively. In the right panel of Fig. 5.3.1 we show the empirical distribution of $\chi_{\rm eff}$ rescaled by the value under the hypotheses that either the secondary or the primary black hole is maximally spinning and aligned with the orbit. We find that several of the events with well-measured nonzero spin do not seem to be well explained by this model (GW190728_064510, GW190521_074359, GW190720_000836, GW170809, GW190930_133541 and GW190828_063405). This is in agreement with earlier findings that either a less extreme model of tidal torques (as argued in (Qin et al., 2018; Bavera et al., 2020)) or a distribution of natal spins with some dispersion is needed in order to explain the observed spins with tides (Roulet et al., 2020).

5.4 Model selection

In order to validate and quantify our findings in Section 5.3, in this section we perform selection of parametric models for the observed binary black hole population. We first provide a brief review of our statistical framework, and then constrain the parameters of several models for the astrophysical effective spin and primary mass distributions.

5.4.1 Statistical framework

Following Roulet et al. (2020), we evaluate the likelihood $P(\{d_i\} \mid \lambda)$ of an observed set of triggers $\{d_i\}$, given a phenomenological population model λ for the distributions of binary black hole source

parameters, as:

$$P(\{d_i\} \mid \lambda) \propto e^{-N_a(\lambda)} \prod_{i=1}^{N_{\rm trig}} \left(\frac{\mathrm{d}N_a(\lambda)}{\mathrm{d}N_a(\lambda_0)}\bigg|_{d_i} p_{{\rm astro},i}(\lambda_0) + 1 - p_{{\rm astro},i}(\lambda_0) \right). \tag{5.4}$$

Here, $N_a(\lambda)$ is the expected number of triggers of astrophysical origin under the population model λ (as opposed to detector noise), over a fixed and arbitrary detection threshold; $\mathrm{d}N_a(\lambda)/\mathrm{d}N_a(\lambda_0)|_{d_i}$ is the ratio of expected densities, in data space, of astrophysical triggers similar to that of the ith event between the population model λ and a fixed, arbitrary reference model λ_0; and $p_{{\rm astro},i}(\lambda_0)$ is the probability of astrophysical origin of the ith trigger under the reference population model. The data space contains observable quantities that carry information about the astrophysical population, like measured detector strains and derived detection statistics. All the quantities described above depend on the search pipeline used; in addition, $N_a(\lambda)$ and the set of triggers itself depend on the detection threshold chosen. Three ingredients are required in order to estimate these quantities: a set of software injections labeled by whether they exceed the detection threshold, to quantify the sensitivity of the search; posterior samples characterizing the parameters of each individual event; and the set of $\{p_{{\rm astro},i}(\lambda_0)\}$ encoding the events' significance (Roulet et al., 2020).

Equation (5.4) naturally factors into the product of likelihoods from searches on disjoint datasets, such as different observing runs. Since the full strain data from observing runs O1 and O2 are publicly accessible (Abbott et al., 2021), for these data sets we base our analysis on our searches for binary black holes (Venumadhav et al., 2019, 2020; Zackay et al., 2019a). The strain data for O3a has also recently been released, and analyzed by Nitz et al. (2021) when this work was close to completion; we do not include these results here. The LVC provides a set of software injections with FAR estimates from the search pipelines they used (cWB, GstLAL, PyCBC and PyCBC BBH), and the GWTC-2 catalog itself which reports $\{p_{{\rm astro},i}\}$ for the latter 3 pipelines (Abbott et al., 2020e). For O3a we use these data products, which are adequate for computing the quantities in Eq. (5.4) with the following caveat. Our method requires knowing $\{p_{{\rm astro},i}(\lambda_0)\}$ under some specific astrophysical model, which was not explicited in the GWTC-2 release. We take two alternative approaches: (i) we conservatively consider only O3a events that are in the gold sample, so that all $p_{\rm astro} = 1$ under any model allowed by observations; or (ii) we consider the same O3a binary black hole mergers as in Ref. (Abbott et al., 2020f)—i.e. with an inverse false-alarm rate IFAR $> 1\,$yr in any pipeline and excluding GW190814—taking the reported $p_{\rm astro}$ at face value and assigning it to an arbitrary model λ_0 featuring a broad distribution in black hole parameters, described below. We will refer to these two samples as GWTC-1 + IAS + Gold O3a and GWTC-1 + IAS + GWTC-2, respectively.

We will find that our conclusions are not strongly affected by the sample used. We implement the sample choices by setting appropriate thresholds on the IFAR, which are reported for both events and injections in GWTC-2. The O3a injections do not report whether they fall near a glitch (one of the criteria of the gold sample), but these should be present in only a few percent of the events given the reported rate of $\sim 1\,\mathrm{glitch/min}$ (Abbott et al., 2020e).

Following (Roulet et al., 2020), we adopt a fiducial population model λ_0 that is described by the following distribution function:

$$f(m_{1\mathrm{s}}, q, \chi_{\mathrm{eff}}, D_L \mid \lambda_0') \propto m_{1\mathrm{s}}^{-\alpha_0} D_L^2, \tag{5.5}$$

where D_L is the luminosity distance and $\alpha_0 = 2.35$. We adopted the λ' notation for the parameters that control the shape of the distribution, while the rate R controls its normalization, i.e. $\lambda = (R, \lambda')$. The ranges of the parameters in Eq. (5.5) are taken to be $m_{1\mathrm{min}} < m_{1\mathrm{s}} < m_{1\mathrm{max}}$ and $q_{\mathrm{min}} < q < 1$, where $m_{1\mathrm{min}} = 3\,\mathrm{M}_\odot$, $m_{1\mathrm{max}} = 120\,\mathrm{M}_\odot$ and $q_{\mathrm{min}} = 1/20$.

5.4.2 Spin distribution

Motivated by Fig. 5.3.1 and the discussion in Section 5.3, as well as Refs. (Farr et al., 2018; Abbott et al., 2020f), we will consider a phenomenological model for the effective spin distribution that allows us to explore the degree of symmetry of the distribution about $\chi_{\mathrm{eff}} = 0$. This model will also allow us to quantify the support for positive and negative values of χ_{eff} in the population.

Firstly, we model the effective spin distribution as a mixture of three subpopulations with negative, zero, and positive χ_{eff}:

$$\begin{aligned} f_{\chi_{\mathrm{eff}}}(\chi_{\mathrm{eff}} \mid \zeta_{\mathrm{pos}}, \zeta_{\mathrm{neg}}, \sigma_{\chi_{\mathrm{eff}}}) = \zeta_0 \mathcal{N}(\chi_{\mathrm{eff}}; \sigma_0 = 0.04) \\ + \zeta_{\mathrm{neg}} \mathcal{N}_{<0}(\chi_{\mathrm{eff}}; \sigma_{\chi_{\mathrm{eff}}}) \\ + \zeta_{\mathrm{pos}} \mathcal{N}_{>0}(\chi_{\mathrm{eff}}; \sigma_{\chi_{\mathrm{eff}}})\,. \end{aligned} \tag{5.6}$$

Here, the various parameters $\zeta_j \in [0, 1]$ are the branching ratios for each subpopulation, constrained to have unit sum; $\mathcal{N}(x; \sigma)$ is the normal distribution with zero mean, dispersion σ, truncated at $x = \pm 1$; $\mathcal{N}_{<0}(x; \sigma)$ is a similar normal distribution but truncated at $x = -1$ and $x = 0$, while $\mathcal{N}_{>0}(x; \sigma)$ is truncated at $x = 0$ and $x = 1$. The functional form Eq. (5.6) is sketched in Fig. 5.4.1 for a particular choice of parameters. Note that we have enforced the dispersion parameters of the positive and negative subpopulations to be equal, such that setting $\zeta_{\mathrm{pos}} = \zeta_{\mathrm{neg}}$ yields a symmetric

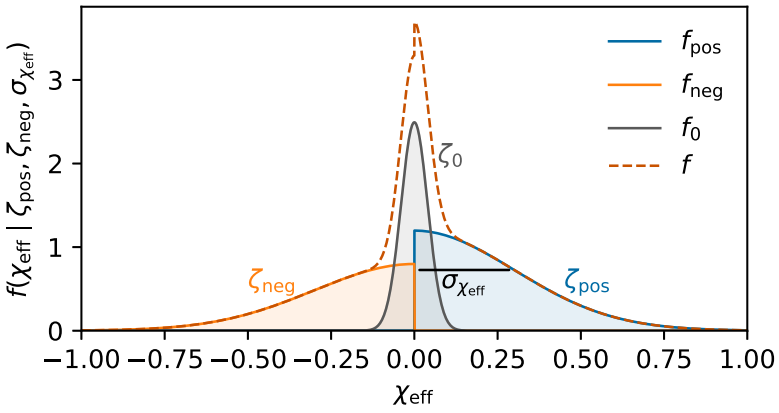


Figure 5.4.1: Sketch of the functional form Eq. (5.6), which we use to parametrize the $\chi_{\rm eff}$ distribution as the sum of three subpopulations. These subpopulations have positive, negative or zero effective spins, with each described by truncated Gaussians that peak at $\chi_{\rm eff} = 0$. We use three independent shape parameters: the width of the positive and negative distributions, $\sigma_{\chi_{\rm eff}}$, which are constrained to be equal, and the three branching ratios ζ_j which sum to unity. For technical reasons, we fix the width of the $\chi_{\rm eff} \approx 0$ subpopulation to have a small but non-vanishing dispersion of $\sigma_0 = 0.04$.

$\chi_{\rm eff}$ distribution. For the $\chi_{\rm eff} \approx 0$ subpopulation, we adopt a small (relative to typical measurement uncertainties) but nonvanishing dispersion $\sigma_0 = 0.04$ in order to ensure that the reweighting procedure used in our algorithm is well behaved (Roulet et al., 2020).

In this Section we will only vary the effective spin distribution, while the remaining spin components are assumed to follow the parameter estimation prior described in Section 5.2. For the other binary black hole parameters, we will assume the following fixed distribution :

$$f(\chi_{\rm eff}, m_{\rm 1s}, q, D_L) = f_{\chi_{\rm eff}}(\chi_{\rm eff}) f_{m_{\rm 1s}}(m_{\rm 1s}) f_q(q) f_{D_L}(D_L). \tag{5.7}$$

Following Abbott et al. (2020f) we adopt a broken power-law distribution for the primary mass:

$$f_{m_{\rm 1s}}(m_{\rm 1s}) \propto \begin{cases} 0, & m_{\rm 1s} < 5\,{\rm M}_\odot \\ \left(\dfrac{m_{\rm 1s}}{m_{\rm break}}\right)^{-\alpha_1}, & 5\,{\rm M}_\odot < m_{\rm 1s} < m_{\rm break} \\ \left(\dfrac{m_{\rm 1s}}{m_{\rm break}}\right)^{-\alpha_2}, & m_{\rm break} < m_{\rm 1s}, \end{cases} \tag{5.8}$$

with $\alpha_1 = 1.6$, $\alpha_2 = 5.6$, $m_{\rm break} = 40\,{\rm M}_\odot$. For simplicity, we adopt a mass-ratio distribution that is uniform in $1/20 < q < 1$ and take the distance distribution to be uniform in comoving volume-time.

We use the likelihood in Eq. (5.4) to obtain a posterior distribution for the population parameters, by adopting a Jeffreys prior for the overall merger rate $\pi(R \mid \lambda') \propto \sqrt{N_a(R, \lambda')}/R$; recall that λ'

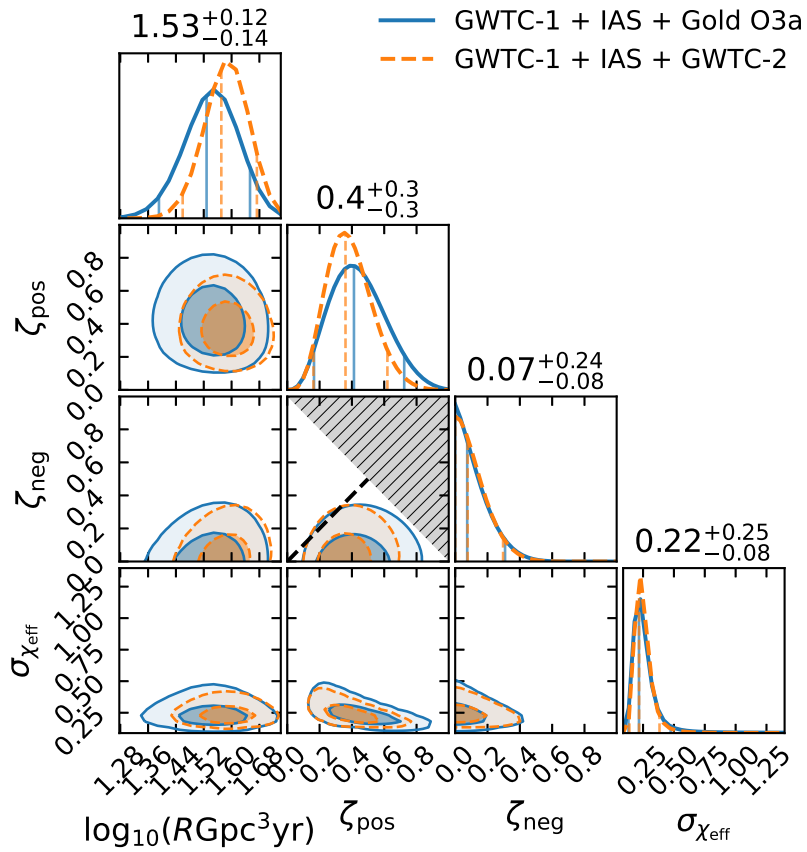


Figure 5.4.2: Constraints on the model parameters of the population model Eq. (5.6). We see that a symmetric $\chi_{\rm eff}$ distribution (black dashed line), with $\zeta_{\rm pos} = \zeta_{\rm neg}$, is disfavored by the data. In addition, the population is consistent with having no negative-spin subpopulation. The two-dimensional contours enclose the 50% and 90% credible regions. Parameter values (median and 90% confidence level) are reported for the GWTC-1 + IAS + Gold O3a sample.

	Δ max ln L	Δ ln Z
Symmetric $\chi_{\rm eff}$	0	0
Positive $\chi_{\rm eff}$	$2.1^{+0.5}_{-0.4}$	$1.6^{+0.5}_{-0.3}$
Positive/Negative mixture $\chi_{\rm eff}$	$2.1^{+0.5}_{-0.4}$	$1.4^{+0.4}_{-0.2}$
Gaussian $\chi_{\rm eff}$	$0.2^{+0.7}_{-0.6}$	$-0.2^{+0.6}_{-0.8}$

Table 5.4.1: Scores for models of the $\chi_{\rm eff}$ distribution. Difference in the maximum log likelihood and log evidence relative to the symmetric $\chi_{\rm eff}$ model $\zeta_{\rm pos} = \zeta_{\rm neg}$. Error bars indicate the 90% confidence level and account for stochastic errors due to the finite number of injections and parameter samples used, and are estimated with 100 bootstrap realizations of the analysis similarly to (Roulet et al., 2020).

are the shape parameters $(\zeta_{\rm pos}, \zeta_{\rm neg}, \sigma_{\chi_{\rm eff}})$. For these we adopt a uniform prior $\pi(\lambda') = {\rm const}$. This prior is invariant to the choice of which two out of the three branching ratios are used to parametrize the distribution.

We show our constraints on the parameters of this model in Fig. 5.4.2, for the two samples used. We find two remarkable results: first, 95% of the posterior lies at $\zeta_{\rm pos} > \zeta_{\rm neg}$ and a symmetric distribution $\zeta_{\rm pos} = \zeta_{\rm neg}$ (dashed line) is disfavored; second, the population is consistent with $\zeta_{\rm neg} = 0$, i.e. no spins anti-aligned with the binary orbit. These conclusions do not depend on which of the two event samples are considered.

We quantify these statements using the Bayesian evidence and maximum likelihood as model scores: we report in Table 5.4.1 the scores achieved by the following models: a symmetric $\chi_{\rm eff}$ distribution given by Eq. (5.6) with $\zeta_{\rm pos} = \zeta_{\rm neg}$, a positive $\chi_{\rm eff}$ distribution setting $\zeta_{\rm neg} = 0$, and the full mixture. The symmetric $\chi_{\rm eff}$ model is representative of a scenario completely dominated by dynamical formation in clusters, while the positive $\chi_{\rm eff}$ model represents a case dominated by isolated binaries—with the caveat that in this channel there exist mechanisms to achieve some spin–orbit misalignment, e.g. supernova kicks.

The first result that positive $\chi_{\rm eff}$ predominates over negative is in general agreement with the analysis of Abbott et al. (2020f). Indeed, parametrizing the $\chi_{\rm eff}$ distribution with a Gaussian, they find that a positive mean is preferred; likewise, they favor spin orientation distributions with at least some degree of anisotropy.

On the other hand, our second finding that there is yet no evidence for negative $\chi_{\rm eff}$ in the population is in contrast with the results of Abbott et al. (2020f), who found that all Gaussian fits to the observed $\chi_{\rm eff}$ distribution had a sizable support at negative $\chi_{\rm eff}$. We suggest that their result is contingent on the assumed parametrization of the population as a Gaussian distribution, while our parametrization has more freedom to accommodate features near $\chi_{\rm eff} = 0$. In particular, the

maximum likelihood solution has parameters $(\zeta_{\rm pos}, \zeta_{\rm neg}, \zeta_0, \sigma_{\chi_{\rm eff}}) = (0.45, 0.00, 0.55, 0.23)$, featuring a sharp peak at $\chi_{\rm eff} \approx 0$, a rapid decline at negative $\chi_{\rm eff}$ and an extended tail at positive $\chi_{\rm eff}$ which are hard to capture with a single Gaussian. To test this hypothesis we try a similar Gaussian model for the $\chi_{\rm eff}$ distribution, shown in Fig. 5.4.3. With this model, we indeed find good quantitative agreement with Abbott et al. (2020f, figure 11) and would recover their same conclusions: we find that models without support at negative $\chi_{\rm eff}$ ($\sigma_{\chi_{\rm eff}} \ll \overline{\chi_{\rm eff}}$) are excluded. In Table 5.4.1 we see that the Gaussian model performs worse than other models we tried, in particular the model restricted to positive $\chi_{\rm eff}$. Abbott et al. (2020f) did consider the possibility that their finding of negative spins could be driven by the Gaussian parametrization. Indeed, in Abbott et al. (2020f, figure 27) they show that adding a free parameter $\chi_{\rm eff}^{\rm min}$ below which the Gaussian is truncated, they exclude $\chi_{\rm eff}^{\rm min} \geq 0$ at 99% credibility and find that small negative values $-0.2 \lesssim \chi_{\rm eff}^{\rm min} \lesssim 0$ are preferred. We interpret that the large number of events at $\chi_{\rm eff} \approx 0$ drives the exclusion of positive $\chi_{\rm eff}^{\rm min}$, furthermore, the fact that small negative values of $\chi_{\rm eff}^{\rm min}$ are preferred over large negative values indicates that the Gaussian model $\chi_{\rm eff}^{\rm min} = -1$, which motivated the claim of existence of negative $\chi_{\rm eff}$ systems, does not fit well the observed population. We conclude that, while it is certainly possible that there are negative $\chi_{\rm eff}$ systems in the population, there is not enough evidence for them yet.

Within isolated formation channels, the fraction of negative $\chi_{\rm eff}$ systems $\zeta_{\rm neg}$ is an indicator of typical natal (supernova) kick velocities, larger kicks generally giving larger $\zeta_{\rm neg}$. Gerosa et al. (2018, figure 6) find that measurements of $\zeta_{\rm neg}$ to a precision better than 0.1 would start putting meaningful constraints on kick velocities. Our current bound $\zeta_{\rm neg} \lesssim 0.3$ is compatible with even extreme kicks, but with a factor of few more detections this would be a promising source of information.

We point out that the GWTC-1 + IAS + GWTC-2 sample differs from that of the analysis in Abbott et al. (2020f) in that it includes events in the IAS catalog. However, having included these events only weakens our conclusions due to the presence of GW170121, the confident detection with the most support for negative $\chi_{\rm eff}$ in the sample.

5.4.3 Mass distribution

We now turn to the distribution of merging binary black hole masses. Using data from the first two observing runs, several past studies have identified that the primary mass distribution was well described by a power-law truncated at $m_{\rm max} \approx 40\,{\rm M}_\odot$ (Fishbach and Holz, 2017; Wysocki et al., 2019; Roulet and Zaldarriaga, 2019; Abbott et al., 2019b; Roulet et al., 2020). The third observing

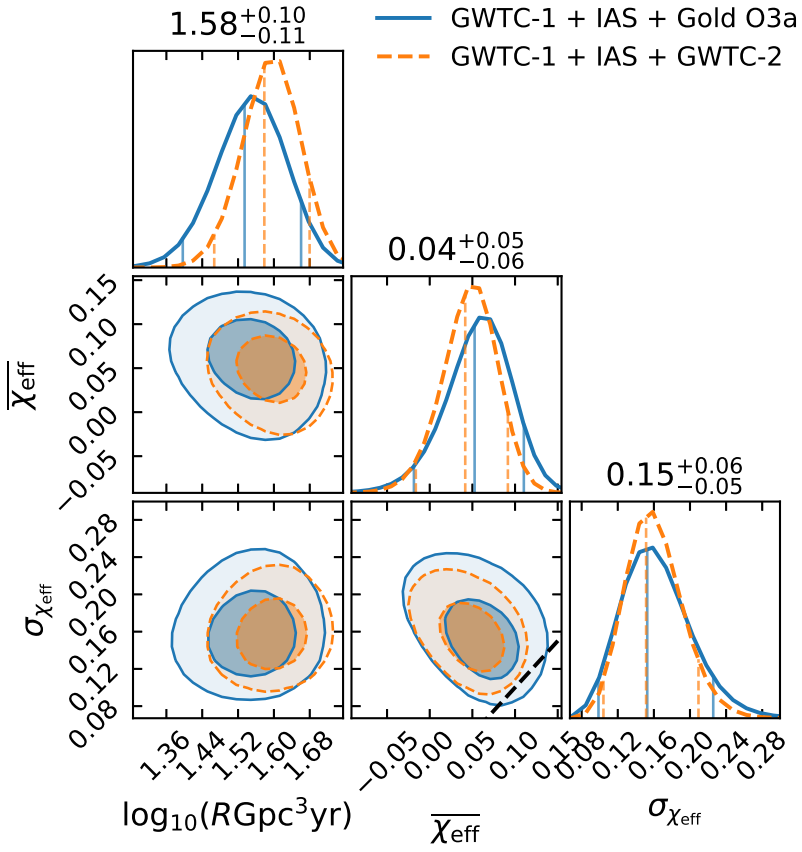


Figure 5.4.3: Constraints assuming a Gaussian model for the $\chi_{\rm eff}$ distribution. The black dashed line corresponds to $\sigma_{\chi_{\rm eff}} = \overline{\chi_{\rm eff}}$, models above the line have sizable support at negative $\chi_{\rm eff}$. Thus, contrary to Fig. 5.4.2, under this model one would conclude that negative $\chi_{\rm eff}$ are present in the population.

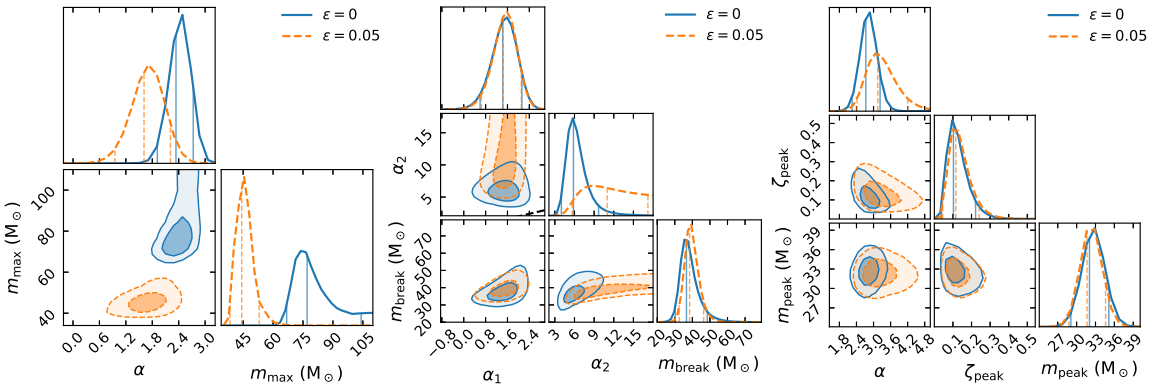
run revealed that the mass distribution has a tail that extends to higher masses, and that models with more features, e.g. a broken power-law, were favored. One diagnostic that a single truncated power-law did not fit the O3a data was that its inferred parameter values experienced a large shift when including the new events, in particular, $m_{\rm max}$ was found to increase from $40.8^{+11.8}_{-4.4}\,\mathrm{M}_\odot$ to $78.5^{+14.1}_{-9.4}\,\mathrm{M}_\odot$ (Abbott et al., 2020f).

As this development evidenced, one has to bear in mind that with a finite number of events one cannot probe the tail of the distribution arbitrarily far out. Thus, constraints obtained on the population are to be interpreted as a characterization of the bulk of the distribution, up to a quantile that depends on the number of events: with $N_{\rm trig}$ triggers, a fraction $\sim \mathcal{O}(1/N_{\rm trig})$ of the distribution cannot be probed; with the present sample this is at the few-percent level. At this point we introduce a feature in our analysis that makes this notion explicit: we add to the model a second subpopulation of astrophysical triggers that come from a broad parameter distribution λ_0' accounting for a small fraction ϵ of the total rate:

$$\frac{\mathrm{d}N_a}{\mathrm{d}\theta}(\theta \mid \lambda, \lambda_0', \epsilon) = R\big[(1-\epsilon) f(\theta \mid \lambda') + \epsilon f(\theta \mid \lambda_0')\big]; \tag{5.9}$$

for $\epsilon = 0$ we recover the previous analysis. Recall that we call the distribution shape parameters λ', so that $\lambda = (R, \lambda')$. For simplicity, we will fix the parameter $\epsilon = 0.05$. This change makes little difference for events that are well described by the population model λ, but since the broad subpopulation can accommodate any of its outliers, the model λ is no longer forced to explain all the observations. A practical advantage of this is that we get a sensitive diagnostic that some specific events may be poorly accommodated by the (ultimately arbitrary) parametrizations we chose, if they get classified with high confidence as belonging to the other subpopulation λ_0'—evidencing that a model with more freedom is needed to explain all events. We also construct a simple goodness-of-fit test for the λ model based on the Bayes factor between a model with $\epsilon = 0$ or a small fixed value $\epsilon = 0.05$. If the $\epsilon = 0$ model is already a good description of all observed events, adding a broad subpopulation should not increase the evidence significantly.

The likelihood for this augmented model can be evaluated in post-processing from the same auxiliary quantities $w_i(\lambda', \lambda_0'), \overline{\mathcal{VT}}(\lambda'), p_{{\rm astro},i}(\lambda_0)$ we use in the evaluation of Eq. (5.4) (see (Roulet



(a) Truncated power law model (b) Broken power law model (c) Power law + peak model

Figure 5.4.4: Adding a broad subpopulation λ_0' with a fraction $\epsilon = 0.05$ of the astrophysical rate affects the inferred parameters of the mass distribution. This is a major effect for the truncated power law model (a), moderate for the broken power law model (b) and minor for the power law + peak model (c). These constraints are derived using the GWTC-1 + IAS + Gold O3a sample of events; we find similar results with GWTC-1 + IAS + GWTC-2.

et al., 2020)):

$$P(\{d_i\} \mid \lambda, \lambda_0', \epsilon) \propto \exp\left\{-R[(1-\epsilon)\overline{\mathcal{VT}}(\lambda') + \epsilon\overline{\mathcal{VT}}(\lambda_0')]\right\} \times \prod_{i=1}^{N_{\rm trig}} \left\{\frac{R}{R_0}\left[w_i(\lambda', \lambda_0')(1-\epsilon) + \epsilon\right] p_{{\rm astro},i}(\lambda_0) + 1 - p_{{\rm astro},i}(\lambda_0)\right\}. \tag{5.10}$$

Likewise, we can also extract the classification of each event as coming from the main component λ or the broader component λ_0': the probability that the ith event came from the λ_0' population is

$$p_{{\rm outlier},i}(\lambda, \lambda_0', \epsilon) = \frac{R\epsilon}{R[(1-\epsilon)w_i(\lambda', \lambda_0') + \epsilon] + R_0(1/p_{{\rm astro},i}(\lambda_0) - 1)}. \tag{5.11}$$

We apply this procedure to three models of the mass distribution that are simplified versions of the Truncated, Broken Power Law and Power Law + Peak models studied in Abbott et al. (2020f). Our broken power law model is given by Eq. (5.8), with $\alpha_1, \alpha_2, m_{\rm break}$ promoted to free parameters. Our truncated model corresponds to $\alpha_2 \to \infty$. Our power law + peak model corresponds to $\alpha_2 = \alpha_1$, plus the addition of a Gaussian component with mean $m_{\rm peak}$ and dispersion $\sigma = 5\,{\rm M}_\odot$ that accounts for a fraction $\zeta_{\rm peak}$ of the total rate. In all three cases we assume a uniform distribution for $\chi_{\rm eff}$, and identical distributions as in Section 5.4.2 for the remaining parameters. With these choices, the models λ' and λ_0' only differ in the primary source-frame mass distribution, which will ease the interpretation of our results.

Fig. 5.4.4 shows the constraints we obtain using the GWTC-1 + IAS + Gold O3a sample; these

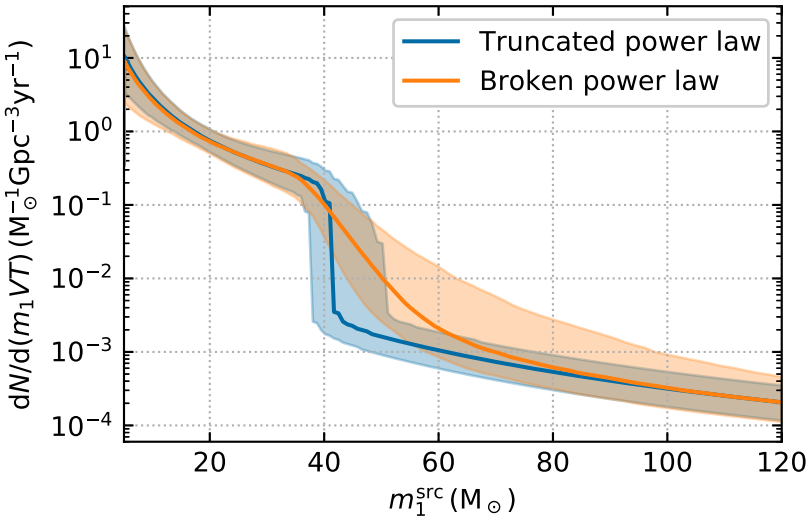


Figure 5.4.5: Mass distribution predicted by the models favored by the data: truncated power law or broken power law, in both cases with an additional broad subpopulation with a fraction $\epsilon = 0.05$ of the total rate responsible for the shallow tail to high masses. These distributions feature a steepening around $40\,M_\odot$, consistent with a step, and a flattening at higher mass. Note that we do not fit for the power-law index nor the normalization of the high-mass tail.

plots are largely unchanged if we use the GWTC-1 + IAS + GWTC-2 sample (not shown). For the case $\epsilon = 0$ we find large quantitative agreement with Abbott et al. (2020f) in the constraints for the corresponding model parameters; in particular, that the data favor a break with $\alpha_1 > \alpha_2$ (region above the dashed line in Fig. 5.4.4b). When we set $\epsilon = 0.05$, allowing these models to not fit all events, we find that the model parameter constraints are affected: for the truncated power-law model the effect is catastrophic, in the sense that the posteriors for $\epsilon = 0$ and 0.05 are inconsistent with each other; while for the broken power-law model there remains a region of overlap and for the power law + peak model the inferred parameters remain largely unaffected. This is in line with the discussion of (Abbott et al., 2020f, figure 2) and suggests that the truncated power-law model with $\epsilon = 0$ fails to describe the astrophysical distribution. It is interesting to note that, with $\epsilon = 0.05$, the truncated and broken power law parametrizations are consistent with the same physical solution $\alpha \approx \alpha_1$, $m_{\rm max} \approx m_{\rm break}$, $\alpha_2 \gg 1$, which exhibits a sharp step at $m_{\rm break}$ and a tail that extends to high masses. The power law + peak parametrization cannot produce a step. We show these inferred distributions in Fig. 5.4.5 to further illustrate the point that both parametrizations give consistent answers, especially for the bulk of the distribution. Note that the differential merger rate is best constrained around $m_{1s} \sim 20\,M_\odot$, where most observations lie (Roulet et al., 2020).

In Table 5.4.2 we report the maximum likelihood and evidence for each of the models studied. We find that, although the broken power law model outperforms the truncated model when $\epsilon = 0$,

	ϵ	Δ max ln L	Δ ln Z
Truncated power law	0	$-7.4^{+0.3}_{-0.3}$	$-6.21^{+0.29}_{-0.19}$
	0.05	0	0
Broken power law	0	$-2.51^{+0.14}_{-0.12}$	$-3.12^{+0.13}_{-0.14}$
	0.05	$-0.23^{+0.07}_{-0.11}$	$0.02^{+0.03}_{-0.04}$
Power law + peak	0	$-1.03^{+0.18}_{-0.17}$	$-3.18^{+0.22}_{-0.13}$
	0.05	$0.04^{+0.21}_{-0.17}$	$-1.56^{+0.15}_{-0.17}$

Table 5.4.2: Scores for models of the primary mass distribution. Maximum log likelihood and log evidence for truncated power law and broken power law models, plus a fraction $\epsilon = 0$ or 0.05 of the population coming from a broad distribution λ_0' per Eq. (5.9). The scores are referred to the preferred truncated power law model with $\epsilon = 0.05$. In all three cases $\epsilon = 0$ is disfavored, implying that the models struggle to accommodate all observations.

both perform poorly relative to their $\epsilon = 0.05$ counterparts. This suggests that neither is a good description of the mass distribution. The power law + peak model achieves similar scores as the broken power law model with $\epsilon = 0$, but it gets only a slight improvement from $\epsilon = 0.05$, thus getting similarly disfavored. Among all the variations, thus, the preferred models in terms of evidence are either the truncated or broken power law with $\epsilon = 0.05$, i.e. with a small fraction of events in a broad tail that extends to high masses. The fact that these two models achieve similar likelihood and evidence, together with the above observation that they are consistent with the same physical solution, suggests that both are comparably good descriptions of the bulk of the distribution and their different scores for $\epsilon = 0$ are driven by the few outlier events. This is confirmed in Fig. 5.4.5. Comparing the $\epsilon = 0$ entries in Table 5.4.1 to Abbott et al. (2020f, table 2), we find agreement in that the truncated power law model is rejected, however, Abbott et al. (2020f) find a preference for the power law + peak model over the broken power law, which we instead find comparable. Some differences are expected because, for simplicity, in our implementation of these models we fixed or omitted some parameters, so the models and associated phase spaces are not equivalent.

We can get some insight by inspecting the probabilities p_{outlier} of coming from the broad subpopulation λ_0' assigned to each event, which we report in Table 5.4.3. Events with a high value of p_{outlier} are better explained by the broad subpopulation and drive a preference for $\epsilon \neq 0$. However, note that even if the true astrophysical population was well described by the parametrization λ, in a catalog of many events some are bound to be in the tail of the distribution and might individually be better described by a broad distribution. The expected distribution of p_{outlier} under a model λ is hard to compute, which is why we do not use the values of p_{outlier} as a quantitative model test. This said, it is apparent that GW190521 is an extreme outlier of the truncated power law model, and there are two other events that are in some tension. For the broken power law model, GW190521

	Truncated power law	Broken power law	Power law + peak
GW190521	1.00	0.94	0.68
GW190602_175927	0.95	0.72	0.66
GW190706_222641	0.88	0.72	0.75
GW190519_153544	0.76	0.54	0.59
GW190929_012149	0.57	0.46	0.51
GW190620_030421	0.43	0.34	0.47
GW190701_203306	0.33	0.19	0.29
GW190413_134308	0.27	0.25	0.31

Table 5.4.3: Probability that each event is a model outlier, as defined in Eq. (5.11) and marginalized over model parameters λ, with $\epsilon = 0.05$, for the mass models studied. Only events in the gold sample with the highest values of $p_{\rm outlier}$ are shown, for brevity. Note that this naturally selects the events with highest primary mass.

is in some tension but the other values of $p_{\rm outlier}$ are milder. For the power law + peak model, no single event is a strong outlier.

Another interesting effect is that GW170817A, a candidate event with $m_{1s} = 56^{+16}_{-10}\,\mathrm{M}_\odot$ and a rather low false-alarm rate of $1/(36\,\mathrm{O2})$ observing runs (Zackay et al., 2019a), had an estimated probability of astrophysical origin marginalized over population parameters of $\overline{p_{\rm astro}} = 0.07$, under the truncated power law model favored after O1 and O2 (Roulet et al., 2020). This low value was driven by the lack of observations of other events with similar properties, mainly mass. Under the newly favored models, it has a moderately different $\overline{p_{\rm astro}} = 0.22$ for the truncated power law and 0.26 for the broken power law, both with $\epsilon = 0.05$. This showcases that $p_{\rm astro}$ values for marginal events in the tails of the distribution are bound to get updated as our knowledge of the population improves.

To summarize, Table 5.4.2, Figs. 5.4.4 and 5.4.5 suggest that the mass distribution exhibits a steepening around $40\,\mathrm{M}_\odot$ and an extended, shallow high-mass tail. From Table 5.4.3 we conclude that the need for this tail is dominated by GW190521, so at this point we do not attempt to model its shape based on a single event. Future data releases will allow to probe these features in the mass distribution.

5.5 Conclusions

We have investigated the properties of the effective spin and primary mass distributions of binary black holes identified in the GWTC-1 (Abbott et al., 2019a), GWTC-2 (Abbott et al., 2020e), and IAS O1–O2 (Venumadhav et al., 2019; Zackay et al., 2019b; Venumadhav et al., 2020; Zackay et al.,

2019a) event catalogs. Our study involved re-analyzing all binary black hole signals with the recently developed `IMRPhenomXPHM` waveform model (Pratten et al., 2020), which includes orbital precession and higher-order modes.

We designed a parametric model of the $\chi_{\rm eff}$ distribution which has three components — with negative, approximately zero and positive $\chi_{\rm eff}$ — to test some general predictions of the dynamic and isolated formation channels for merging binary black holes. Namely, dynamical formation channels predict a $\chi_{\rm eff}$ distribution that is symmetric about 0, while negative $\chi_{\rm eff}$ (i.e., large spin–orbit misalignment) should be very rare for isolated field binaries. Interestingly, we found that a symmetric distribution is disfavored: the data suggests that the number of positive $\chi_{\rm eff}$ events is larger than that with negative $\chi_{\rm eff}$ at 95% credibility. Although the evidence at this point is not conclusive, this simple test is already becoming powerful enough to hint that not all binary black holes are dynamically assembled, in agreement with other analyses of these data (Abbott et al., 2020f; Zevin et al., 2021; Bouffanais et al., 2021). The number of detections is expected to roughly double with the forthcoming release of the O3b catalog, which should settle this question if the same trend continues.

Moreover, we find no evidence for negative $\chi_{\rm eff}$ in the population. This result is in tension with Ref. (Abbott et al., 2020f); we attribute the discrepancy to the different parametrizations of the spin distribution chosen. We were able to reproduce the results of Ref. (Abbott et al., 2020f) with a Gaussian model for $\chi_{\rm eff}$, but found that this model fares worse at describing the features in the spin distribution, in particular, a large concentration of events near $\chi_{\rm eff} = 0$. Our conclusion is in agreement with a model-free inspection of the empirical $\chi_{\rm eff}$ distribution, which suggests that all events with significant support at $\chi_{\rm eff} < 0$ are consistent with coming from a population with $\chi_{\rm eff} = 0$. Therefore, we conclude that the observed effective spin distribution does not rule out that all observations are explained by isolated binary formation.

Regarding the distribution of primary masses, we confirmed the result of (Abbott et al., 2020f) that a truncated power law fails to describe the observations. Moreover, we found evidence that a broken power law model or a power law plus a Gaussian peak, which assume a continuous distribution, compare poorly to a model in which a small fraction of the events comes from a broad subpopulation, with an extended tail at high masses. This suggests that the tail of the mass distribution has interesting features that will be probed with the coming data releases.

Acknowledgements

We are grateful to Ilya Mandel, Will Farr, Davide Gerosa and the attendees of the gravitational wave meeting at the Center for Computational Astrophysics for helpful comments and discussion.

HSC gratefully acknowledges support from the Rubicon Fellowship awarded by the Netherlands Organisation for Scientific Research (NWO). SO acknowledges support from the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-2039656. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. LD acknowledges support from the Michael M. Garland startup research grant at the University of California, Berkeley. TV acknowledges support by the National Science Foundation under Grant No. 2012086. BZ is supported by a research grant from the Ruth and Herman Albert Scholarship Program for New Scientists. MZ is supported by NSF grants PHY-1820775 the Canadian Institute for Advanced Research (CIFAR) Program on Gravity and the Extreme Universe and the Simons Foundation Modern Inflationary Cosmology initiative.

This research has made use of data, software and/or web tools obtained from the Gravitational Wave Open Science Center (`https://www.gw-openscience.org/`), a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. LIGO Laboratory and Advanced LIGO are funded by the United States National Science Foundation (NSF) as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. Virgo is funded, through the European Gravitational Observatory (EGO), by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale di Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, Spain.

Appendix

5.A Sample Selection

In this appendix, we inventorize the binary black hole mergers used in this work, which are listed in Table 5.A.1. We define the gold sample (third column of Table 5.A.1) as the set of events that (i) were detected by at least two search pipelines with a FAR $< 0.1\,\mathrm{yr}^{-1}$ (fourth column); and (ii) on strain data that are free of non-Gaussian transient noise (fifth column). We consider the following pipelines: cWB (Abbott et al., 2020e), GstLAL (Abbott et al., 2020e), PyCBC (Abbott et al., 2020e; Nitz et al., 2019a, 2020), PyCBC BBH (Abbott et al., 2020e; Nitz et al., 2020), and IAS (Venumadhav et al., 2020). Events with non-Gaussian artifacts are reported in (Abbott et al., 2020f, table V). We do not include GW190814 in the GWTC-1 + IAS + Gold O3a sample because it was detected near non-Gaussian transient noise (Abbott et al., 2020e). Nor do we include GW190814 in the GWTC-1 + IAS + GWTC-2 sample (Section 5.4.2) as it was not included in the main GWTC-2 population analysis due to being an outlier in the mass ratio distribution (Abbott et al., 2020f). For events in the O1 and O2 observing runs, $p_{\mathrm{astro}}(\lambda_0)$ is computed in (Roulet et al., 2020). For events in O3a, it is taken at face value from (Abbott et al., 2020e) as the maximum p_{astro} over pipelines, and may not accurately correspond to the model λ_0.

While the present work was being completed, Nitz et al. (2021) reported their analysis of the O3a data, providing independent confirmation of all the sources reported in GWTC-2 except for GW190426_152155 and GW190909_114149, and further finding four previously unreported events. We defer the inclusion of these results to future work. Including this catalog, the two-pipeline condition would be fulfilled by most of the O3a events in Table 5.A.1, thereby enlarging the gold sample. Still, note that the sample restriction did not change the qualitative conclusions of our analysis.

Run	Name	Gold	≥ 2 pip.	Clean	$p_{\mathrm{astro}}(\lambda_0)$
O1	GW150914	✓	✓	✓	1.00
	GW151012	✓	✓	✓	1.00
	GW151226	✓	✓	✓	1.00
	GW151216			✓	0.50
O2	GW170823	✓	✓	✓	1.00
	GW170809	✓	✓	✓	1.00
	GW170729	✓	✓	✓	1.00
	GW170814	✓	✓	✓	1.00
	GW170104	✓	✓	✓	1.00
	GW170727	✓	✓	✓	0.99
	GW170121	✓	✓	✓	0.97
	GW170304			✓	1.00
	GW170818	✓	✓	✓	0.92
	170412B			✓	0.02
	GW170403			✓	0.61
	GW170425			✓	0.60
	GW170202			✓	0.61
	GW170817A			✓	0.74
	GW170608	✓	✓	✓	1.00
O3a	GW190408_181802	✓	✓	✓	1.00
	GW190412	✓	✓	✓	1.00
	GW190413_052954			✓	0.98
	GW190413_134308				0.98
	GW190421_213856	✓	✓	✓	1.00
	GW190424_180648				0.91
	GW190503_185404		✓		1.00
	GW190512_180714	✓	✓	✓	1.00
	GW190513_205428		✓		1.00
	GW190514_065416				0.96
	GW190517_055101	✓	✓	✓	1.00
	GW190519_153544	✓	✓	✓	1.00
	GW190521	✓	✓	✓	1.00
	GW190521_074359	✓	✓	✓	1.00
	GW190527_092055			✓	0.99
	GW190602_175927	✓	✓	✓	1.00
	GW190620_030421			✓	1.00
	GW190630_185205			✓	1.00
	GW190701_203306				1.00
	GW190706_222641	✓	✓	✓	1.00
	GW190707_093326	✓	✓	✓	1.00
	GW190708_232457			✓	1.00
	GW190719_215514			✓	0.82
	GW190720_000836	✓	✓	✓	1.00
	GW190727_060333		✓		1.00
	GW190728_064510	✓	✓	✓	1.00
	GW190731_140936			✓	0.97
	GW190803_022701	✓	✓	✓	0.99
	GW190828_063405	✓	✓	✓	1.00
	GW190828_065509	✓	✓	✓	1.00
	GW190909_114149			✓	0.89
	GW190910_112807			✓	1.00
	GW190915_235702	✓	✓	✓	1.00
	GW190924_021846		✓		1.00
	GW190929_012149			✓	1.00
	GW190930_133541	✓	✓	✓	1.00

Table 5.A.1: Binary black hole events used in this work. Checkmarks from the third to fifth columns indicate events that are in the gold sample, were identified by at least two pipelines with IFAR $>$ 10 yr, and was observed in the absence of glitches, respectively. The p_{astro} values shown here are evaluated with the reference model λ_0 described in Eq. (5.5).

Chapter 6

Conclusion

In this thesis we made the journey from raw gravitational wave data to its implications for the astrophysics of binary black holes. In the process we developed a search pipeline with several innovations, including a geometric template bank. We approximately doubled the sensitive volume of previous searches, allowing us to find nine new events in LIGO–Virgo O1 and O2 public datasets. We introduced and implemented a statistical framework to extract the astrophysical interpretation of these signals, which required careful accounting of their measurement uncertainty, selection effects and statistical significance. Over these few years, we saw the field of gravitational wave astrophysics quickly evolve as new data was gathered. The picture is improving, but still far from clear for binary black hole formation channels. Some key takeaway points that we have learned are that the distribution of black hole spins is anisotropic, with a preference for alignment with the orbital angular momentum. This disfavors dynamical assembly of binary black holes in dense clusters as the only formation channel in operation. On the other hand, the black hole mass distribution features a tail to high masses inside the mass-gap region predicted by the pair instability supernova process, which is challenging for binaries formed in isolation but a natural prediction of dynamical channels. Regarding the merger rate, we found that it is much better constrained if one restricts it to the region of parameter space where most detections lie. We expect this realization to enable powerful tests of physically-motivated binary black hole population synthesis predictions.

This thesis highlights the importance of public data, without which none of these projects could have taken place. My collaborators and I are very grateful to the LIGO–Virgo Collaboration for their open data policy.

Looking forward, we are quickly reaching the point where improvements in detector sensitivity

and gravitational waveform models enable measurement of subtle subdominant General Relativity effects that carry crucial astrophysical information. Inspiraling binaries emit dominantly quadrupolar gravitational radiation, yet very recently higher-order modes have been unambiguously detected. These enable new tests of General Relativity and break parameter degeneracies, allowing precise measurement of the component masses and distance to the source. Spin–orbit precession occurs only if the black hole spins and binary orbit are misaligned, which in turn would shed light on the formation mechanism of the system: misaligned spins are natural for binaries formed dynamically from independent black holes in dense stellar environments, but are harder to produce for BBHs formed from isolated binary stars. Likewise, detecting an eccentric merger would clearly indicate a prompt dynamical formation, as binary orbits circularize on a shorter timescale than they merge. Despite the motivations, state-of-the-art search pipelines do not include these physical effects, presenting a direct avenue for improving our sensitivity to these most interesting sources.

Coupled to the steady progress in the waveform modeling front, parameter estimation methods are another area of active development. Computational cost and sensitivity to algorithm systematics are currently a significant bottleneck in analysis workflows. The forecasted dramatic increase in rate of detections, together with the need of using more advanced and computationally expensive waveform models (and testing them on software injections to calibrate systematic errors) will require major advancements in parameter estimation algorithms.

Altogether, it is clear that the field of gravitational wave astrophysics is undergoing rapid progress and that the coming few years are going to be full of delightful surprises.

Bibliography

J. Aasi et al. Advanced LIGO. *Classical and Quantum Gravity*, 32(7):074001, Mar. 2015. doi: 10.1088/0264-9381/32/7/074001. URL https://doi.org/10.1088/0264-9381/32/7/074001.

B. P. Abbott et al. Observation of gravitational waves from a binary black hole merger. *Physical Review Letters*, 116:061102, 2016a. doi: 10.1103/PhysRevLett.116.061102. URL https://link.aps.org/doi/10.1103/PhysRevLett.116.061102.

B. P. Abbott et al. GW151226: Observation of gravitational waves from a 22-solar-mass binary black hole coalescence. *Physical Review Letters*, 116:241103, 2016b. doi: 10.1103/PhysRevLett.116.241103. URL https://link.aps.org/doi/10.1103/PhysRevLett.116.241103.

B. P. Abbott et al. Binary black hole mergers in the first Advanced LIGO observing run. *Phys. Rev. X*, 6:041015, Oct 2016c. doi: 10.1103/PhysRevX.6.041015. URL https://link.aps.org/doi/10.1103/PhysRevX.6.041015.

B. P. Abbott et al. The basic physics of the binary black hole merger GW150914. *Annalen der Physik*, 529(1-2):1600209, oct 2016d. doi: 10.1002/andp.201600209. URL https://doi.org/10.1002/andp.201600209.

B. P. Abbott et al. Search for intermediate mass black hole binaries in the first observing run of Advanced LIGO. *Physical Review*, D96(2):022001, 2017a. doi: 10.1103/PhysRevD.96.022001.

B. P. Abbott et al. GW170104: Observation of a 50-solar-mass binary black hole coalescence at redshift 0.2. *Phys. Rev. Lett.*, 118:221101, Jun 2017b. doi: 10.1103/PhysRevLett.118.221101. URL https://link.aps.org/doi/10.1103/PhysRevLett.118.221101.

B. P. Abbott et al. GW170608: Observation of a 19 solar-mass binary black hole coalescence. *The Astrophysical Journal*, 851(2):L35, 2017c. doi: 10.3847/2041-8213/aa9f0c. URL https://doi.org/10.3847/2041-8213/aa9f0c.

B. P. Abbott et al. GW170814: A three-detector observation of gravitational waves from a binary black hole coalescence. *Physical Review Letters*, 119:141101, 2017d. doi: 10.1103/PhysRevLett.119.141101. URL https://link.aps.org/doi/10.1103/PhysRevLett.119.141101.

B. P. Abbott et al. GW170817: Observation of gravitational waves from a binary neutron star inspiral. *Phys. Rev. Lett.*, 119:161101, Oct 2017e. doi: 10.1103/PhysRevLett.119.161101. URL https://link.aps.org/doi/10.1103/PhysRevLett.119.161101.

B. P. Abbott et al. GWTC-1: A gravitational-wave transient catalog of compact binary mergers observed by LIGO and Virgo during the first and second observing runs. *Phys. Rev. X*, 9: 031040, Sep 2019a. doi: 10.1103/PhysRevX.9.031040. URL https://link.aps.org/doi/10.1103/PhysRevX.9.031040.

B. P. Abbott et al. Binary black hole population properties inferred from the first and second observing runs of Advanced LIGO and Advanced Virgo. *The Astrophysical Journal*, 882(2): L24, Sept. 2019b. doi: 10.3847/2041-8213/ab3800. URL https://doi.org/10.3847/2041-8213/ab3800.

B. P. Abbott et al. GW190425: Observation of a compact binary coalescence with total mass $\sim 3.4\,\mathrm{M}_\odot$. *The Astrophysical Journal*, 892(1):L3, Mar. 2020a. doi: 10.3847/2041-8213/ab75f5. URL https://doi.org/10.3847/2041-8213/ab75f5.

R. Abbott et al. GW190412: Observation of a binary-black-hole coalescence with asymmetric masses. *Phys. Rev. D*, 102:043015, Aug 2020b. doi: 10.1103/PhysRevD.102.043015. URL https://link.aps.org/doi/10.1103/PhysRevD.102.043015.

R. Abbott et al. GW190521: A binary black hole merger with a total mass of $150\,\mathrm{M}_\odot$. *Physical Review Letters*, 125(10), Sept. 2020c. doi: 10.1103/physrevlett.125.101102. URL https://doi.org/10.1103/physrevlett.125.101102.

R. Abbott et al. GW190814: Gravitational waves from the coalescence of a 23 solar mass black hole with a 2.6 solar mass compact object. *The Astrophysical Journal*, 896(2):L44, June 2020d. doi: 10.3847/2041-8213/ab960f. URL https://doi.org/10.3847/2041-8213/ab960f.

R. Abbott et al. GWTC-2: Compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run, 2020e.

R. Abbott et al. Population properties of compact objects from the second LIGO–Virgo Gravitational-Wave Transient Catalog, 2020f.

R. Abbott et al. Open data from the first and second observing runs of Advanced LIGO and Advanced Virgo. *SoftwareX*, 13:100658, Jan. 2021. doi: 10.1016/j.softx.2021.100658. URL https://doi.org/10.1016/j.softx.2021.100658.

F. Acernese et al. Advanced Virgo: a second-generation interferometric gravitational wave detector. *Classical and Quantum Gravity*, 32(2):024001, Dec. 2014. doi: 10.1088/0264-9381/32/2/024001. URL https://doi.org/10.1088/0264-9381/32/2/024001.

M. Adachi and M. Kasai. An analytical approximation of the luminosity distance in flat cosmologies with a cosmological constant. *Progress of Theoretical Physics*, 127(1):145–152, jan 2012. doi: 10.1143/ptp.127.145. URL https://doi.org/10.1143/ptp.127.145.

P. A. R. Ade et al. Planck2015 results. *Astronomy & Astrophysics*, 594:A13, sep 2016. doi: 10.1051/0004-6361/201525830. URL https://doi.org/10.1051/0004-6361/201525830.

P. Ajith, S. Babak, Y. Chen, M. Hewitson, B. Krishnan, A. M. Sintes, J. T. Whelan, B. Brügmann, P. Diener, N. Dorband, J. Gonzalez, M. Hannam, S. Husa, D. Pollney, L. Rezzolla, L. Santamaría, U. Sperhake, and J. Thornburg. Template bank for gravitational waveforms from coalescing binary black holes: Nonspinning binaries. *Phys. Rev. D*, 77:104017, May 2008. doi: 10.1103/PhysRevD.77.104017. URL https://link.aps.org/doi/10.1103/PhysRevD.77.104017.

P. Ajith, N. Fotopoulos, S. Privitera, A. Neunzert, N. Mazumder, and A. J. Weinstein. Effectual template bank for the detection of gravitational waves from inspiralling compact binaries with generic spins. *Phys. Rev. D*, 89:084041, Apr 2014. doi: 10.1103/PhysRevD.89.084041. URL https://link.aps.org/doi/10.1103/PhysRevD.89.084041.

P. Ajith et al. Inspiral-merger-ringdown waveforms for black-hole binaries with non-precessing spins. *Phys. Rev. Lett.*, 106:241101, 2011. doi: 10.1103/PhysRevLett.106.241101.

B. Allen. χ^2 time-frequency discriminator for gravitational wave detection. *Phys. Rev. D*, 71: 062001, Mar 2005. doi: 10.1103/PhysRevD.71.062001. URL https://link.aps.org/doi/10.1103/PhysRevD.71.062001.

B. Allen, W. G. Anderson, P. R. Brady, D. A. Brown, and J. D. E. Creighton. FINDCHIRP: An algorithm for detection of gravitational waves from inspiraling compact binaries. *Phys. Rev. D*, 85:122006, Jun 2012. doi: 10.1103/PhysRevD.85.122006. URL https://link.aps.org/doi/10.1103/PhysRevD.85.122006.

J. M. Antelis and C. Moreno. An independent search of gravitational waves in the first observation run of Advanced LIGO using cross-correlation. *General Relativity and Gravitation*, 51(5), May 2019. doi: 10.1007/s10714-019-2546-x. URL https://doi.org/10.1007/s10714-019-2546-x.

F. Antonini and H. B. Perets. Secular evolution of compact binaries near massive black holes: Gravitational wave sources and other exotica. *The Astrophysical Journal*, 757(1):27, 2012.

F. Antonini and F. A. Rasio. Merging black hole binaries in galactic nuclei: implications for Advanced-LIGO detections. *The Astrophysical Journal*, 831(2):187, 2016.

F. Antonini, N. Murray, and S. Mikkola. Black hole triple dynamics: a breakdown of the orbit average approximation and implications for gravitational wave detections. *The Astrophysical Journal*, 781 (1):45, 2014.

F. Antonini, S. Toonen, and A. S. Hamers. Binary black hole mergers from field triples: Properties, rates, and the impact of stellar evolution. *The Astrophysical Journal*, 841(2):77, 2017.

T. A. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne. Spin-induced orbital precession and its modulation of the gravitational waveforms from merging binaries. *Physical Review D*, 49: 6274–6297, 1994. doi: 10.1103/PhysRevD.49.6274. URL https://link.aps.org/doi/10.1103/PhysRevD.49.6274.

G. Ashton and E. Thrane. The astrophysical odds of GW151216. *Monthly Notices of the Royal Astronomical Society*, 498(2):1905–1910, 08 2020. ISSN 0035-8711. doi: 10.1093/mnras/staa2332. URL https://doi.org/10.1093/mnras/staa2332.

A. Askar, M. Szkudlarek, D. Gondek-Rosińska, M. Giersz, and T. Bulik. MOCCA-SURVEY database–I. Coalescing binary black holes originating from globular clusters. *Monthly Notices of the Royal Astronomical Society: Letters*, 464(1):L36–L40, 2016.

S. Babak, R. Balasubramanian, D. Churches, T. Cokelaer, and B. S. Sathyaprakash. A template bank to search for gravitational waves from inspiralling compact binaries: I. Physical models. *Classical and Quantum Gravity*, 23(18):5477–5504, aug 2006. doi: 10.1088/0264-9381/23/18/002. URL https://doi.org/10.1088/0264-9381/23/18/002.

S. Babak, R. Biswas, P. R. Brady, D. A. Brown, K. Cannon, C. D. Capano, J. H. Clayton, T. Cokelaer, J. D. E. Creighton, T. Dent, A. Dietz, S. Fairhurst, N. Fotopoulos, G. González, C. Hanna, I. W. Harry, G. Jones, D. Keppel, D. J. A. McKechan, L. Pekowsky, S. Privitera,

C. Robinson, A. C. Rodriguez, B. S. Sathyaprakash, A. S. Sengupta, M. Vallisneri, R. Vaulin, and A. J. Weinstein. Searching for gravitational waves from binary coalescence. *Phys. Rev. D*, 87:024033, Jan 2013. doi: 10.1103/PhysRevD.87.024033. URL `https://link.aps.org/doi/10.1103/PhysRevD.87.024033`.

E. Baird, S. Fairhurst, M. Hannam, and P. Murphy. Degeneracy between mass and spin in black-hole-binary waveforms. *Physical Review D*, 87(2), jan 2013. doi: 10.1103/physrevd.87.024035. URL `https://doi.org/10.1103/physrevd.87.024035`.

S. Banerjee. Stellar-mass black holes in young massive and open stellar clusters and their role in gravitational-wave generation. *Monthly Notices of the Royal Astronomical Society*, 467(1):524–539, 2017.

Z. Barkat, G. Rakavy, and N. Sack. Dynamics of supernova explosion resulting from pair formation. *Phys. Rev. Lett.*, 18:379–381, Mar 1967. doi: 10.1103/PhysRevLett.18.379. URL `https://link.aps.org/doi/10.1103/PhysRevLett.18.379`.

I. Bartos, B. Kocsis, Z. Haiman, and S. Márka. Rapid and bright stellar-mass binary black hole mergers in active galactic nuclei. *The Astrophysical Journal*, 835(2):165, 2017.

S. S. Bavera, T. Fragos, Y. Qin, E. Zapartas, C. J. Neijssel, I. Mandel, A. Batta, S. M. Gaebel, C. Kimball, and S. Stevenson. The origin of spin in binary black holes. *Astronomy & Astrophysics*, 635:A97, Mar. 2020. doi: 10.1051/0004-6361/201936204. URL `https://doi.org/10.1051/0004-6361/201936204`.

K. Belczynski and T. Bulik. The effect of supernova natal kicks on compact object merger rate. *Astronomy and Astrophysics*, 346(1):91–100, 1999.

K. Belczynski, V. Kalogera, and T. Bulik. A comprehensive study of binary compact objects as gravitational wave sources: evolutionary channels, rates, and physical properties. *The Astrophysical Journal*, 572(1):407, 2002.

K. Belczynski, R. E. Taam, V. Kalogera, F. A. Rasio, and T. Bulik. On the rarity of double black hole binaries: Consequences for gravitational wave detection. *The Astrophysical Journal*, 662(1):504, 2007.

K. Belczynski, V. Kalogera, F. A. Rasio, R. E. Taam, A. Zezas, T. Bulik, T. J. Maccarone, and N. Ivanova. Compact object modeling with the startrack population synthesis code. *The Astrophysical Journal Supplement Series*, 174(1):223, 2008.

K. Belczynski, A. Buonanno, M. Cantiello, C. L. Fryer, D. E. Holz, I. Mandel, M. C. Miller, and M. Walczak. The formation and gravitational-wave detection of massive stellar black hole binaries. *The Astrophysical Journal*, 789(2):120, 2014.

K. Belczynski et al. The origin of low spin of black holes in LIGO/Virgo mergers, 2017.

C. M. Biwer, C. D. Capano, S. De, M. Cabero, D. A. Brown, A. H. Nitz, and V. Raymond. PyCBC Inference: A python-based parameter estimation toolkit for compact binary coalescence signals, 2018.

A. Bohé et al. Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with Advanced detectors. *Physical Review D*, 95(4), 2017. doi: 10.1103/physrevd.95.044028. URL `https://doi.org/10.1103/physrevd.95.044028`.

J. R. Bond, W. D. Arnett, and B. J. Carr. The evolution and fate of very massive objects. *The Astrophysical Journal*, 280:825, May 1984. doi: 10.1086/162057. URL `https://doi.org/10.1086/162057`.

Y. Bouffanais, M. Mapelli, F. Santoliquido, N. Giacobbo, U. N. D. Carlo, S. Rastello, M. C. Artale, and G. Iorio. New insights on binary black hole formation channels after GWTC-2: young star clusters versus isolated binaries, 2021.

D. A. Brown, I. Harry, A. Lundgren, and A. H. Nitz. Detecting binary neutron star systems with spin in advanced gravitational-wave detectors. *Phys. Rev. D*, 86:084017, Oct 2012. doi: 10.1103/PhysRevD.86.084017. URL `https://link.aps.org/doi/10.1103/PhysRevD.86.084017`.

J. Buchner, A. Georgakakis, K. Nandra, L. Hsu, C. Rangel, M. Brightman, A. Merloni, M. Salvato, J. Donley, and D. Kocevski. X-ray spectral modelling of the AGN obscuring region in the CDFS: Bayesian model selection and catalogue. *Astronomy & Astrophysics*, 564:A125, Apr. 2014. doi: 10.1051/0004-6361/201322971. URL `https://doi.org/10.1051/0004-6361/201322971`.

T. A. Callister, W. M. Farr, and M. Renzo. State of the fieldmandel2016inary black hole natal kicks and prospects for isolated field formation after GWTC-2, 2020.

M. Campanelli, C. O. Lousto, and Y. Zlochower. Spinning-black-hole binaries: The orbital hang-up. *Physical Review D*, 74(4), 2006. doi: 10.1103/physrevd.74.041501. URL `https://doi.org/10.1103/physrevd.74.041501`.

T. D. Canton and I. W. Harry. Designing a template bank to observe compact binary coalescences in Advanced LIGO's second observing run, 2017.

C. Capano, I. Harry, S. Privitera, and A. Buonanno. Implementing a search for gravitational waves from binary black holes with nonprecessing spin. *Phys. Rev. D*, 93:124007, Jun 2016. doi: 10.1103/PhysRevD.93.124007. URL https://link.aps.org/doi/10.1103/PhysRevD.93.124007.

S. Chatterjee, C. L. Rodriguez, V. Kalogera, and F. A. Rasio. Dynamical formation of low-mass merging black hole binaries like GW151226. *The Astrophysical Journal Letters*, 836(2):L26, 2017.

H.-Y. Chen, R. Essick, S. Vitale, D. E. Holz, and E. Katsavounidis. Observational selection effects with ground-based gravitational wave detectors. *The Astrophysical Journal*, 835(1):31, 2017. doi: 10.3847/1538-4357/835/1/31. URL https://doi.org/10.3847/1538-4357/835/1/31.

H. S. Chia, S. Olsen, J. Roulet, L. Dai, T. Venumadhav, B. Zackay, and M. Zaldarriaga. Boxing day surprise: Higher multipoles and orbital precession in GW151226, 2021.

T. Cokelaer. Gravitational waves from inspiralling compact binaries: Hexagonal template placement and its efficiency in detecting physical signals. *Phys. Rev. D*, 76:102004, Nov 2007. doi: 10.1103/PhysRevD.76.102004. URL https://link.aps.org/doi/10.1103/PhysRevD.76.102004.

N. Cornish, D. Blas, and G. Nardini. Bounding the speed of gravity with gravitational wave observations. *Physical Review Letters*, 119(16), 2017. doi: 10.1103/physrevlett.119.161102. URL https://doi.org/10.1103/physrevlett.119.161102.

N. J. Cornish and T. B. Littenberg. Bayeswave: Bayesian inference for gravitational wave bursts and instrument glitches. *Classical and Quantum Gravity*, 32(13):135012, June 2015. doi: 10.1088/0264-9381/32/13/135012. URL https://doi.org/10.1088/0264-9381/32/13/135012.

N. J. Cornish, T. B. Littenberg, B. Bécsy, K. Chatziioannou, J. A. Clark, S. Ghonge, and M. Millhouse. Bayeswave analysis pipeline in the era of gravitational wave observations. *Phys. Rev. D*, 103:044006, Feb 2021. doi: 10.1103/PhysRevD.103.044006. URL https://link.aps.org/doi/10.1103/PhysRevD.103.044006.

R. Cotesta, A. Buonanno, A. Bohé, A. Taracchini, I. Hinder, and S. Ossokine. Enriching the symphony of gravitational waves from binary black holes by tuning higher harmonics. *Physical Review D*, 98(8), oct 2018. doi: 10.1103/physrevd.98.084028. URL https://doi.org/10.1103/physrevd.98.084028.

C. Cutler and E. E. Flanagan. Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral waveform? *Phys. Rev. D*, 49:2658–2697, Mar 1994. doi: 10.1103/PhysRevD.49.2658. URL https://link.aps.org/doi/10.1103/PhysRevD.49.2658.

L. Dai, B. Zackay, T. Venumadhav, J. Roulet, and M. Zaldarriaga. Search for lensed gravitational waves including Morse phase information: An intriguing candidate in O2, 2020.

S. de Mink and I. Mandel. The chemically homogeneous evolutionary channel for binary black hole mergers: rates and properties of gravitational-wave events detectable by Advanced ligo. *Monthly Notices of the Royal Astronomical Society*, 460(4):3545–3553, 2016.

S. V. Dhurandhar and B. S. Sathyaprakash. Choice of filters for the detection of gravitational waves from coalescing binaries. II. Detection in colored noise. *Phys. Rev. D*, 49:1707–1722, Feb 1994. doi: 10.1103/PhysRevD.49.1707. URL https://link.aps.org/doi/10.1103/PhysRevD.49.1707.

M. Dominik, K. Belczynski, C. Fryer, D. E. Holz, E. Berti, T. Bulik, I. Mandel, and R. O'Shaughnessy. Double compact objects. II. Cosmological merger rates. *The Astrophysical Journal*, 779(1):72, 2013.

J. Downing, M. Benacquista, M. Giersz, and R. Spurzem. Compact binaries in star clusters–I. Black hole binaries inside globular clusters. *Monthly Notices of the Royal Astronomical Society*, 407(3): 1946–1962, 2010.

J. Downing, M. Benacquista, M. Giersz, and R. Spurzem. Compact binaries in star clusters–II. Escapers and detection rates. *Monthly Notices of the Royal Astronomical Society*, 416(1):133–147, 2011.

J. Eldridge and E. Stanway. Bpass predictions for binary black hole mergers. *Monthly Notices of the Royal Astronomical Society*, 462(3):3302–3313, 2016.

H. Estellés, S. Husa, M. Colleoni, M. Mateu-Lucena, M. de Lluc Planas, C. García-Quirós, D. Keitel, A. Ramos-Buades, A. K. Mehta, A. Buonanno, and S. Ossokine. A detailed analysis of GW190521 with phenomenological waveform models, 2021.

R. Farmer, M. Renzo, S. E. de Mink, P. Marchant, and S. Justham. Mind the gap: The location of the lower edge of the pair-instability supernova black hole mass gap. *The Astrophysical Journal*, 887 (1):53, Dec. 2019. doi: 10.3847/1538-4357/ab518b. URL https://doi.org/10.3847/1538-4357/ab518b.

B. Farr, D. E. Holz, and W. M. Farr. Using spin to understand the formation of LIGO and Virgo's black holes. *The Astrophysical Journal*, 854(1):L9, 2018. doi: 10.3847/2041-8213/aaaa64. URL https://doi.org/10.3847/2041-8213/aaaa64.

W. M. Farr, S. Stevenson, M. C. Miller, I. Mandel, B. Farr, and A. Vecchio. Distinguishing spin-aligned and isotropic black hole populations with gravitational waves. *Nature*, 548(7667):426–429, 2017. ISSN 14764687. doi: 10.1038/nature23453.

N. Farrow, X.-J. Zhu, and E. Thrane. The mass distribution of galactic double neutron stars. *The Astrophysical Journal*, 876(1):18, Apr. 2019. doi: 10.3847/1538-4357/ab12e3. URL https://doi.org/10.3847/1538-4357/ab12e3.

M. Fishbach and D. E. Holz. Where are LIGO's big black holes? *The Astrophysical Journal*, 851(2): L25, Dec. 2017. doi: 10.3847/2041-8213/aa9bf6. URL https://doi.org/10.3847/2041-8213/aa9bf6.

M. Fishbach and D. E. Holz. Picky partners: The pairing of component masses in binary black hole mergers. *The Astrophysical Journal*, 891(1):L27, Mar. 2020. doi: 10.3847/2041-8213/ab7247. URL https://doi.org/10.3847/2041-8213/ab7247.

M. Fishbach, D. E. Holz, and W. M. Farr. Does the black hole merger rate evolve with redshift? *The Astrophysical Journal*, 863(2):L41, Aug. 2018. doi: 10.3847/2041-8213/aad800. URL https://doi.org/10.3847/2041-8213/aad800.

M. Fishbach, W. M. Farr, and D. E. Holz. The most massive binary black hole detections and the identification of population outliers. *The Astrophysical Journal*, 891(2):L31, Mar. 2020. doi: 10.3847/2041-8213/ab77c9. URL https://doi.org/10.3847/2041-8213/ab77c9.

W. A. Fowler and F. Hoyle. Neutrino processes and pair formation in massive stars and supernovae. *The Astrophysical Journal Supplement Series*, 9:201, Dec. 1964. doi: 10.1086/190103. URL https://doi.org/10.1086/190103.

J. Fuller and L. Ma. Most black holes are born very slowly rotating. *The Astrophysical Journal*, 881(1):L1, Aug. 2019. doi: 10.3847/2041-8213/ab339b. URL https://doi.org/10.3847/2041-8213/ab339b.

B. Gadre, S. Mitra, and S. Dhurandhar. Hierarchical search strategy for the efficient detection of gravitational waves from nonprecessing coalescing compact binaries with aligned-spins. *Phys. Rev.*

D, 99:124035, Jun 2019. doi: 10.1103/PhysRevD.99.124035. URL https://link.aps.org/doi/10.1103/PhysRevD.99.124035.

S. M. Gaebel, J. Veitch, T. Dent, and W. M. Farr. Digging the population of compact binary mergers out of the noise. *Monthly Notices of the Royal Astronomical Society*, 484(3):4008–4023, Jan. 2019. doi: 10.1093/mnras/stz225. URL https://doi.org/10.1093/mnras/stz225.

S. Galaudage, C. Talbot, and E. Thrane. Gravitational-wave inference in the catalog era: Evolving priors and marginal events. *Phys. Rev. D*, 102:083026, Oct 2020. doi: 10.1103/PhysRevD.102.083026. URL https://link.aps.org/doi/10.1103/PhysRevD.102.083026.

V. Gayathri, I. Bartos, Z. Haiman, S. Klimenko, B. Kocsis, S. Márka, and Y. Yang. GW170817A as a hierarchical black hole merger. *The Astrophysical Journal*, 890(2):L20, Feb. 2020. doi: 10.3847/2041-8213/ab745d. URL https://doi.org/10.3847/2041-8213/ab745d.

D. Gerosa and M. Fishbach. Hierarchical mergers of stellar-mass black holes and their gravitational-wave signatures, 2021.

D. Gerosa, M. Kesden, E. Berti, R. O'Shaughnessy, and U. Sperhake. Resonant-plane locking and spin alignment in stellar-mass black-hole binaries: A diagnostic of compact-binary formation. *Phys. Rev. D*, 87:104028, May 2013. doi: 10.1103/PhysRevD.87.104028. URL https://link.aps.org/doi/10.1103/PhysRevD.87.104028.

D. Gerosa, E. Berti, R. O'Shaughnessy, K. Belczynski, M. Kesden, D. Wysocki, and W. Gladysz. Spin orientations of merging black holes formed from the evolution of stellar binaries. *Physical Review D*, 98(8), oct 2018. doi: 10.1103/physrevd.98.084036. URL https://doi.org/10.1103/physrevd.98.084036.

N. Giacobbo and M. Mapelli. The progenitors of compact-object binaries: impact of metallicity, common envelope and natal kicks. *Monthly Notices of the Royal Astronomical Society*, 480(2):2011–2030, July 2018. doi: 10.1093/mnras/sty1999. URL https://doi.org/10.1093/mnras/sty1999.

N. Giacobbo, M. Mapelli, and M. Spera. Merging black hole binaries: the effects of progenitor's metallicity, mass-loss rate and Eddington factor. *Monthly Notices of the Royal Astronomical Society*, 474(3):2959–2974, 2017.

A. S. Hamers, H. B. Perets, F. Antonini, and S. F. P. Zwart. Secular dynamics of hierarchical quadruple systems: the case of a triple system orbited by a fourth body. *Monthly Notices of*

the Royal Astronomical Society, 449(4):4221–4245, Apr. 2015. doi: 10.1093/mnras/stv452. URL https://doi.org/10.1093/mnras/stv452.

I. W. Harry, B. Allen, and B. S. Sathyaprakash. Stochastic template placement algorithm for gravitational wave data analysis. *Phys. Rev. D*, 80:104014, Nov 2009. doi: 10.1103/PhysRevD.80.104014. URL https://link.aps.org/doi/10.1103/PhysRevD.80.104014.

I. W. Harry, A. H. Nitz, D. A. Brown, A. P. Lundgren, E. Ochsner, and D. Keppel. Investigating the effect of precession on searches for neutron-star–black-hole binaries with Advanced LIGO. *Phys. Rev. D*, 89:024010, Jan 2014. doi: 10.1103/PhysRevD.89.024010. URL https://link.aps.org/doi/10.1103/PhysRevD.89.024010.

A. Heger, C. L. Fryer, S. E. Woosley, N. Langer, and D. H. Hartmann. How massive single stars end their life. *The Astrophysical Journal*, 591(1):288–300, July 2003. doi: 10.1086/375341. URL https://doi.org/10.1086/375341.

K. Hotekezaka and T. Piran. Are the observed black hole mergers spins consistent with field binary progenitors?, 2017.

K. Hotokezaka and T. Piran. Implications of the low binary black hole aligned spins observed by LIGO. *The Astrophysical Journal*, 842(2):111, June 2017. doi: 10.3847/1538-4357/aa6f61. URL https://doi.org/10.3847/1538-4357/aa6f61.

Y. Huang, C.-J. Haster, S. Vitale, A. Zimmerman, J. Roulet, T. Venumadhav, B. Zackay, L. Dai, and M. Zaldarriaga. Source properties of the lowest signal-to-noise-ratio binary black hole detections, 2020.

P. Jaranowski and A. Królak. Gravitational-wave data analysis. Formalism and sample applications: The Gaussian case. *Living Reviews in Relativity*, 15(1):4, Mar 2012. ISSN 1433-8351. doi: 10.12942/lrr-2012-4. URL https://doi.org/10.12942/lrr-2012-4.

S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé. Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era. *Phys. Rev. D*, 93:044007, Feb 2016. doi: 10.1103/PhysRevD.93.044007. URL https://link.aps.org/doi/10.1103/PhysRevD.93.044007.

L. E. Kidder. Coalescing binary systems of compact objects to post-Newtonian 5/2 order. 5. Spin effects. *Phys. Rev. D*, 52:821–847, 1995. doi: 10.1103/PhysRevD.52.821.

T. O. Kimpson, M. Spera, M. Mapelli, and B. M. Ziosi. Hierarchical black hole triples in young star clusters: impact of kozai–lidov resonance on mergers. *Monthly Notices of the Royal Astronomical Society*, 463(3):2443–2452, 2016.

S. Klimenko, G. Vedovato, M. Drago, F. Salemi, V. Tiwari, G. A. Prodi, C. Lazzaro, K. Ackley, S. Tiwari, C. F. Da Silva, and G. Mitselmakher. Method for detection and reconstruction of gravitational wave transients with networks of advanced detectors. *Phys. Rev. D*, 93:042004, Feb 2016. doi: 10.1103/PhysRevD.93.042004. URL https://link.aps.org/doi/10.1103/PhysRevD.93.042004.

M. U. Kruckow, T. M. Tauris, N. Langer, M. Kramer, and R. G. Izzard. Progenitors of gravitational wave mergers: Binary evolution with the stellar grid-based code ComBinE. *Monthly Notices of the Royal Astronomical Society*, 481(2):1908–1949, Aug. 2018. doi: 10.1093/mnras/sty2190. URL https://doi.org/10.1093/mnras/sty2190.

D. Kushnir, M. Zaldarriaga, J. A. Kollmeier, and R. Waldman. GW150914: Spin-based constraints on the merger time of the progenitor system. *Monthly Notices of the Royal Astronomical Society*, 462(1):844–849, 07 2016. ISSN 0035-8711. doi: 10.1093/mnras/stw1684. URL https://doi.org/10.1093/mnras/stw1684.

T. B. Littenberg, J. B. Kanner, N. J. Cornish, and M. Millhouse. Enabling high confidence detections of gravitational-wave bursts. *Phys. Rev. D*, 94:044050, Aug 2016. doi: 10.1103/PhysRevD.94.044050. URL https://link.aps.org/doi/10.1103/PhysRevD.94.044050.

B. Liu and D. Lai. Black hole and neutron star binary mergers in triple systems: Merger fraction and spin–orbit misalignment. *The Astrophysical Journal*, 863(1):68, Aug. 2018. doi: 10.3847/1538-4357/aad09f. URL https://doi.org/10.3847/1538-4357/aad09f.

K.-W. Lo and L.-M. Lin. The spin parameter of uniformly rotating compact stars. *The Astrophysical Journal*, 728(1):12, Jan. 2011. doi: 10.1088/0004-637x/728/1/12. URL https://doi.org/10.1088/0004-637x/728/1/12.

T. J. Loredo. Accounting for source uncertainties in analyses of astronomical survey data. In *AIP Conference Proceedings*. AIP, 2004. doi: 10.1063/1.1835214. URL https://doi.org/10.1063/1.1835214.

R. Lynch, M. Coughlin, S. Vitale, C. W. Stubbs, and E. Katsavounidis. Observational implications

of lowering the LIGO-Virgo alert threshold. *The Astrophysical Journal*, 861(2):L24, jul 2018. doi: 10.3847/2041-8213/aacf9f. URL https://doi.org/10.3847/2041-8213/aacf9f.

R. Magee, H. Fong, S. Caudill, C. Messick, K. Cannon, P. Godwin, C. Hanna, S. Kapadia, D. Meacher, S. R. Mohite, D. Mukherjee, A. Pace, S. Sachdev, M. Shikauchi, and L. Singer. Sub-threshold binary neutron star search in Advanced LIGO's first observing run. *The Astrophysical Journal*, 878(1):L17, jun 2019. doi: 10.3847/2041-8213/ab20cf. URL https://doi.org/10.3847%2F2041-8213%2Fab20cf.

I. Mandel and S. E. de Mink. Merging binary black holes formed through chemically homogeneous evolution in short-period stellar binaries. *Monthly Notices of the Royal Astronomical Society*, 458 (3):2634–2647, feb 2016. doi: 10.1093/mnras/stw379. URL https://doi.org/10.1093/mnras/stw379.

I. Mandel and A. Farmer. Merging stellar-mass binary black holes, 2018.

I. Mandel and R. O'Shaughnessy. Compact binary coalescences in the band of ground-based gravitational-wave detectors. *Classical and Quantum Gravity*, 27(11):114007, may 2010. doi: 10.1088/0264-9381/27/11/114007. URL https://doi.org/10.1088/0264-9381/27/11/114007.

I. Mandel, W. M. Farr, and J. R. Gair. Extracting distribution parameters from multiple uncertain observations with selection biases. *Monthly Notices of the Royal Astronomical Society*, 486(1): 1086–1093, Mar. 2019. doi: 10.1093/mnras/stz896. URL https://doi.org/10.1093/mnras/stz896.

M. Mapelli. Massive black hole binaries from runaway collisions: the impact of metallicity. *Monthly Notices of the Royal Astronomical Society*, 459(4):3432–3446, 2016.

M. Mapelli and N. Giacobbo. The cosmic merger rate of neutron stars and black holes. *Monthly Notices of the Royal Astronomical Society*, 2018.

M. Mapelli, N. Giacobbo, E. Ripamonti, and M. Spera. The cosmic merger rate of stellar black hole binaries from the illustris simulation. *Monthly Notices of the Royal Astronomical Society*, 472(2): 2422–2435, 2017.

P. Marchant, N. Langer, P. Podsiadlowski, T. M. Tauris, and T. J. Moriya. A new route towards merging massive black holes. *Astronomy & Astrophysics*, 588:A50, 2016.

J. E. McClintock, R. Narayan, and J. F. Steiner. Black hole spin via continuum fitting and the role of spin in powering transient jets. *Space Science Reviews*, 183(1-4):295–322, 2013. doi: 10.1007/s11214-013-0003-9. URL https://doi.org/10.1007/s11214-013-0003-9.

B. McKernan, K. Ford, W. Lyra, and H. Perets. Intermediate mass black holes in agn discs–i. production and growth. *Monthly Notices of the Royal Astronomical Society*, 425(1):460–469, 2012.

N. Mennekens and D. Vanbeveren. Massive double compact object mergers: gravitational wave sources and r-process element production sites. *Astronomy & Astrophysics*, 564:A134, 2014.

M. C. Miller and J. M. Miller. The masses and spins of neutron stars and stellar-mass black holes. *Physics Reports*, 548:1–34, Jan. 2015. doi: 10.1016/j.physrep.2014.09.003. URL https://doi.org/10.1016/j.physrep.2014.09.003.

S. Miller, T. A. Callister, and W. M. Farr. The low effective spin of binary black holes and implications for individual gravitational-wave events. *The Astrophysical Journal*, 895(2):128, June 2020. doi: 10.3847/1538-4357/ab80c0. URL https://doi.org/10.3847/1538-4357/ab80c0.

G. Nelemans, L. Yungelson, and S. P. Zwart. The gravitational wave signal from the galactic disk population of binaries containing two compact objects. *Astronomy & Astrophysics*, 375(3):890–898, 2001.

K. K. Y. Ng, S. Vitale, A. Zimmerman, K. Chatziioannou, D. Gerosa, and C.-J. Haster. Gravitational-wave astrophysics with effective-spin measurements: Asymmetries and selection biases. *Phys. Rev. D*, 98:083007, Oct 2018. doi: 10.1103/PhysRevD.98.083007. URL https://link.aps.org/doi/10.1103/PhysRevD.98.083007.

A. H. Nitz and C. D. Capano. GW190521 may be an intermediate-mass ratio inspiral. *The Astrophysical Journal*, 907(1):L9, Jan. 2021. doi: 10.3847/2041-8213/abccc5. URL https://doi.org/10.3847/2041-8213/abccc5.

A. H. Nitz, T. Dent, T. D. Canton, S. Fairhurst, and D. A. Brown. Detecting binary compact-object mergers with gravitational waves: Understanding and improving the sensitivity of the PyCBC search. *The Astrophysical Journal*, 849(2):118, Nov. 2017. doi: 10.3847/1538-4357/aa8f50. URL https://doi.org/10.3847/1538-4357/aa8f50.

A. H. Nitz, C. Capano, A. B. Nielsen, S. Reyes, R. White, D. A. Brown, and B. Krishnan. 1-OGC: The first open gravitational-wave catalog of binary mergers from analysis of public Advanced

LIGO data. *The Astrophysical Journal*, 872(2):195, Feb. 2019a. doi: 10.3847/1538-4357/ab0108. URL https://doi.org/10.3847/1538-4357/ab0108.

A. H. Nitz, A. B. Nielsen, and C. D. Capano. Potential gravitational-wave and gamma-ray multi-messenger candidate from 2015 October 30. *The Astrophysical Journal*, 876(1):L4, Apr. 2019b. doi: 10.3847/2041-8213/ab18a1. URL https://doi.org/10.3847/2041-8213/ab18a1.

A. H. Nitz, T. Dent, G. S. Davies, S. Kumar, C. D. Capano, I. Harry, S. Mozzon, L. Nuttall, A. Lundgren, and M. Tápai. 2-OGC: Open gravitational-wave catalog of binary mergers from analysis of public Advanced LIGO and Virgo data. *The Astrophysical Journal*, 891(2):123, Mar. 2020. doi: 10.3847/1538-4357/ab733f. URL https://doi.org/10.3847/1538-4357/ab733f.

A. H. Nitz, C. D. Capano, S. Kumar, Y.-F. Wang, S. Kastha, M. Schäfer, R. Dhurkunde, and M. Cabero. 3-OGC: Catalog of gravitational waves from compact-binary mergers, 2021.

S. R. Olsen et al. in preparation.

B. J. Owen. Search templates for gravitational waves from inspiraling binaries: Choice of template spacing. *Phys. Rev. D*, 53:6749–6761, Jun 1996. doi: 10.1103/PhysRevD.53.6749. URL https://link.aps.org/doi/10.1103/PhysRevD.53.6749.

B. J. Owen and B. S. Sathyaprakash. Matched filtering of gravitational waves from inspiraling compact binaries: Computational cost and template placement. *Phys. Rev. D*, 60:022002, Jun 1999. doi: 10.1103/PhysRevD.60.022002. URL https://link.aps.org/doi/10.1103/PhysRevD.60.022002.

R. M. O'Leary, F. A. Rasio, J. M. Fregeau, N. Ivanova, and R. O'Shaughnessy. Binary mergers and growth of black holes in dense star clusters. *The Astrophysical Journal*, 637(2):937, 2006.

C. Petrovich and F. Antonini. Greatly enhanced merger rates of compact-object binaries in non-spherical nuclear star clusters. *The Astrophysical Journal*, 846(2):146, 2017.

Z. Piran and T. Piran. The origin of binary black hole mergers. *The Astrophysical Journal*, 892(1):64, Mar. 2020. doi: 10.3847/1538-4357/ab792a. URL https://doi.org/10.3847/1538-4357/ab792a.

G. Pratten and A. Vecchio. Assessing gravitational-wave binary black hole candidates with Bayesian odds, 2020.

G. Pratten, C. García-Quirós, M. Colleoni, A. Ramos-Buades, H. Estellés, M. Mateu-Lucena, R. Jaume, M. Haney, D. Keitel, J. E. Thompson, and S. Husa. Computationally efficient models for the dominant and sub-dominant harmonic modes of precessing binary black holes, 2020.

S. Privitera, S. R. P. Mohapatra, P. Ajith, K. Cannon, N. Fotopoulos, M. A. Frei, C. Hanna, A. J. Weinstein, and J. T. Whelan. Improving the sensitivity of a search for coalescing binary black holes with nonprecessing spins in gravitational wave data. *Phys. Rev. D*, 89:024003, Jan 2014. doi: 10.1103/PhysRevD.89.024003. URL https://link.aps.org/doi/10.1103/PhysRevD.89.024003.

M. Pürrer, M. Hannam, and F. Ohme. Can we measure individual black-hole spins from gravitational-wave observations? *Phys. Rev. D*, 93(8):084042, 2016. doi: 10.1103/PhysRevD.93.084042.

Y. Qin, T. Fragos, G. Meynet, J. Andrews, M. Sørensen, and H. F. Song. The spin of the second-born black hole in coalescing binary black holes. *Astronomy & Astrophysics*, 616:A28, aug 2018. doi: 10.1051/0004-6361/201832839. URL https://doi.org/10.1051/0004-6361/201832839.

É. Racine. Analysis of spin precession in binary black hole systems including quadrupole-monopole interaction. *Physical Review D*, 78(4), aug 2008. doi: 10.1103/physrevd.78.044021. URL https://doi.org/10.1103/physrevd.78.044021.

C. L. Rodriguez and F. Antonini. A triple origin for the heavy and low-spin binary black holes detected by LIGO/Virgo. *The Astrophysical Journal*, 863(1):7, aug 2018. doi: 10.3847/1538-4357/aacea4. URL https://doi.org/10.3847/1538-4357/aacea4.

C. L. Rodriguez, M. Morscher, B. Pattabiraman, S. Chatterjee, C.-J. Haster, and F. A. Rasio. Binary black hole mergers from globular clusters: Implications for Advanced LIGO. *Phys. Rev. Lett.*, 115:051101, Jul 2015. doi: 10.1103/PhysRevLett.115.051101. URL https://link.aps.org/doi/10.1103/PhysRevLett.115.051101.

C. L. Rodriguez, S. Chatterjee, and F. A. Rasio. Binary black hole mergers from globular clusters: Masses, merger rates, and the impact of stellar evolution. *Physical Review D*, 93(8):084029, 2016a.

C. L. Rodriguez, M. Zevin, C. Pankow, V. Kalogera, and F. A. Rasio. Illuminating black hole binary formation channels with spins in Advanced LIGO. *The Astrophysical Journal*, 832(1):L2, nov 2016b. doi: 10.3847/2041-8205/832/1/l2. URL https://doi.org/10.3847/2041-8205/832/1/l2.

C. L. Rodriguez, P. Amaro-Seoane, S. Chatterjee, and F. A. Rasio. Post-Newtonian dynamics in dense star clusters: Highly eccentric, highly spinning, and repeated binary black hole mergers. *Physical Review Letters*, 120(15), apr 2018. doi: 10.1103/physrevlett.120.151101. URL https://doi.org/10.1103/physrevlett.120.151101.

J. Roulet and M. Zaldarriaga. Constraints on binary black hole populations from LIGO–Virgo detections. *Monthly Notices of the Royal Astronomical Society*, 484(3):4216–4229, jan 2019. doi: 10.1093/mnras/stz226. URL https://doi.org/10.1093/mnras/stz226.

J. Roulet, L. Dai, T. Venumadhav, B. Zackay, and M. Zaldarriaga. Template bank for compact binary coalescence searches in gravitational wave data: A general geometric placement algorithm. *Phys. Rev. D*, 99:123022, Jun 2019. doi: 10.1103/PhysRevD.99.123022. URL https://link.aps.org/doi/10.1103/PhysRevD.99.123022.

J. Roulet, T. Venumadhav, B. Zackay, L. Dai, and M. Zaldarriaga. Binary black hole mergers from LIGO/Virgo O1 and O2: Population inference combining confident and marginal events. *Phys. Rev. D*, 102:123022, Dec 2020. doi: 10.1103/PhysRevD.102.123022. URL https://link.aps.org/doi/10.1103/PhysRevD.102.123022.

J. Roulet, H. S. Chia, S. Olsen, L. Dai, T. Venumadhav, B. Zackay, and M. Zaldarriaga. On the distribution of effective spins and masses of binary black holes from the ligo and virgo o1-o3a observing runs, 2021.

S. Roy, A. S. Sengupta, and N. Thakor. Hybrid geometric-random template-placement algorithm for gravitational wave searches from compact binary coalescences. *Phys. Rev. D*, 95:104045, May 2017. doi: 10.1103/PhysRevD.95.104045. URL https://link.aps.org/doi/10.1103/PhysRevD.95.104045.

S. Roy, A. S. Sengupta, and P. Ajith. Effectual template banks for upcoming compact binary searches in Advanced-LIGO and Virgo data. *Phys. Rev. D*, 99:024048, Jan 2019. doi: 10.1103/PhysRevD.99.024048. URL https://link.aps.org/doi/10.1103/PhysRevD.99.024048.

S. Sachdev, S. Caudill, H. Fong, R. K. L. Lo, C. Messick, D. Mukherjee, R. Magee, L. Tsukada, K. Blackburn, P. Brady, P. Brockill, K. Cannon, S. J. Chamberlin, D. Chatterjee, J. D. E. Creighton, P. Godwin, A. Gupta, C. Hanna, S. Kapadia, R. N. Lang, T. G. F. Li, D. Meacher, A. Pace, S. Privitera, L. Sadeghian, L. Wade, M. Wade, A. Weinstein, and S. L. Xiao. The

GstLAL search analysis methods for compact binary mergers in Advanced LIGO's second and Advanced Virgo's first observing runs, 2019.

A. Sadowski, K. Belczynski, T. Bulik, N. Ivanova, F. A. Rasio, and R. O'Shaughnessy. The total merger rate of compact object binaries in the local universe. *The Astrophysical Journal*, 676(2): 1162, 2008.

M. Safarzadeh, W. M. Farr, and E. Ramirez-Ruiz. A trend in the effective spin distribution of LIGO binary black holes with mass. *The Astrophysical Journal*, 894(2):129, May 2020. doi: 10.3847/1538-4357/ab80be. URL `https://doi.org/10.3847/1538-4357/ab80be`.

J. Samsing, M. MacLeod, and E. Ramirez-Ruiz. The Formation of Eccentric Compact Binary Inspirals and the Role of Gravitational Wave Emission in Binary-Single Stellar Encounters. *Astrophys. J.*, 784:71, 2014. doi: 10.1088/0004-637X/784/1/71.

L. Santamaria et al. Matching post-Newtonian and numerical relativity waveforms: systematic errors and a new phenomenological model for non-precessing black hole binaries. *Phys. Rev. D*, 82:064016, 2010. doi: 10.1103/PhysRevD.82.064016.

B. S. Sathyaprakash and B. F. Schutz. Physics, astrophysics and cosmology with gravitational waves. *Living Reviews in Relativity*, 12(1), mar 2009. doi: 10.12942/lrr-2009-2. URL `https://doi.org/10.12942/lrr-2009-2`.

P. R. Saulson. "Sufficiently advanced technology" for gravitational wave detection, 2019.

B. F. Schutz. Networks of gravitational wave detectors and three figures of merit. *Classical and Quantum Gravity*, 28(12):125023, May 2011. doi: 10.1088/0264-9381/28/12/125023. URL `https://doi.org/10.1088/0264-9381/28/12/125023`.

R. J. E. Smith, C. Talbot, F. H. Vivanco, and E. Thrane. Inferring the population properties of binary black holes from unresolved gravitational waves. *Monthly Notices of the Royal Astronomical Society*, 496(3):3281–3290, June 2020. doi: 10.1093/mnras/staa1642. URL `https://doi.org/10.1093/mnras/staa1642`.

M. Spera, M. Mapelli, and A. Bressan. The mass spectrum of compact remnants from the parsec stellar evolution tracks. *Monthly Notices of the Royal Astronomical Society*, 451(4):4086–4103, 2015.

N. Steinle and M. Kesden. Pathways for producing binary black holes with large misaligned spins in the isolated formation channel, 2020.

S. Stevenson, C. P. L. Berry, and I. Mandel. Hierarchical analysis of gravitational-wave measurements of binary black hole spin–orbit misalignments. *Monthly Notices of the Royal Astronomical Society*, 471(3):2801–2811, jul 2017a. doi: 10.1093/mnras/stx1764. URL https://doi.org/10.1093/mnras/stx1764.

S. Stevenson, A. Vigna-Gómez, I. Mandel, J. W. Barrett, C. J. Neijssel, D. Perkins, and S. E. de Mink. Formation of the first three gravitational-wave observations through isolated binary evolution. *Nature communications*, 8:14906, 2017b.

N. C. Stone, B. D. Metzger, and Z. Haiman. Assisted inspirals of stellar mass black holes embedded in agn discs: solving the 'final au problem'. *Monthly Notices of the Royal Astronomical Society*, 464(1):946–954, 2016.

N. Tacik, F. Foucart, H. P. Pfeiffer, R. Haas, S. Ossokine, J. Kaplan, C. Muhlberger, M. D. Duez, L. E. Kidder, M. A. Scheel, and B. Szilágyi. Binary neutron stars with arbitrary spins in numerical relativity. *Phys. Rev. D*, 92:124012, Dec 2015. doi: 10.1103/PhysRevD.92.124012. URL https://link.aps.org/doi/10.1103/PhysRevD.92.124012.

C. Talbot and E. Thrane. Determining the population properties of spinning black holes. *Phys. Rev. D*, 96:023012, Jul 2017. doi: 10.1103/PhysRevD.96.023012. URL https://link.aps.org/doi/10.1103/PhysRevD.96.023012.

T. Tanaka and H. Tagoshi. Use of new coordinates for the template space in a hierarchical search for gravitational waves from inspiraling binaries. *Phys. Rev. D*, 62:082001, Sep 2000. doi: 10.1103/PhysRevD.62.082001. URL https://link.aps.org/doi/10.1103/PhysRevD.62.082001.

S. R. Taylor and D. Gerosa. Mining gravitational-wave catalogs to understand binary stellar evolution: A new hierarchical Bayesian framework. *Physical Review D*, 98(8), oct 2018. doi: 10.1103/physrevd.98.083017. URL https://doi.org/10.1103/physrevd.98.083017.

K. S. Thorne. *Gravitational radiation.*, pages 330–458. Cambridge University Press, 1987.

E. Thrane and C. Talbot. An introduction to Bayesian inference in gravitational-wave astronomy: Parameter estimation, model selection, and hierarchical models. *Publications of the Astronomical Society of Australia*, 36, 2019. doi: 10.1017/pasa.2019.2. URL https://doi.org/10.1017/pasa.2019.2.

V. Tiwari. Estimation of the sensitive volume for gravitational-wave source populations using weighted monte carlo integration. *Classical and Quantum Gravity*, 35(14):145009, June 2018. doi: 10.1088/1361-6382/aac89d. URL https://doi.org/10.1088/1361-6382/aac89d.

S. A. Usman, A. H. Nitz, I. W. Harry, C. M. Biwer, D. A. Brown, M. Cabero, C. D. Capano, T. Dal Canton, T. Dent, S. Fairhurst, M. S. Kehl, D. Keppel, B. Krishnan, A. Lenon, A. Lundgren, A. B. Nielsen, L. P. Pekowsky, H. P. Pfeiffer, P. R. Saulson, M. West, and J. L. Willis. The PyCBC search for gravitational waves from compact binary coalescence. *Classical and Quantum Gravity*, 33:215004, Nov 2016. doi: 10.1088/0264-9381/33/21/215004.

M. Vallisneri, J. Kanner, R. Williams, A. Weinstein, and B. Stephens. The LIGO Open Science Center. *Journal of Physics: Conference Series*, 610:012021, 2015. doi: 10.1088/1742-6596/610/1/012021. URL https://doi.org/10.1088/1742-6596/610/1/012021.

T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga. New search pipeline for compact binary mergers: Results for binary black holes in the first observing run of Advanced LIGO. *Phys. Rev. D*, 100:023011, Jul 2019. doi: 10.1103/PhysRevD.100.023011. URL https://link.aps.org/doi/10.1103/PhysRevD.100.023011.

T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga. New binary black hole mergers in the second observing run of Advanced LIGO and Advanced Virgo. *Phys. Rev. D*, 101:083030, Apr 2020. doi: 10.1103/PhysRevD.101.083030. URL https://link.aps.org/doi/10.1103/PhysRevD.101.083030.

S. Vitale. One, no one, and one hundred thousand – inferring the properties of a population in presence of selection effects, 2020.

S. Vitale, R. Lynch, J. Veitch, V. Raymond, and R. Sturani. Measuring the spin of black holes in binary systems using gravitational waves. *Physical Review Letters*, 112(25), 2014. doi: 10.1103/physrevlett.112.251101. URL https://doi.org/10.1103/physrevlett.112.251101.

S. Vitale, R. Lynch, V. Raymond, R. Sturani, J. Veitch, and P. Graff. Parameter estimation for heavy binary-black holes with networks of second-generation gravitational-wave detectors. *Physical Review D*, 95(6), 2017a. doi: 10.1103/physrevd.95.064053. URL https://doi.org/10.1103/physrevd.95.064053.

S. Vitale, R. Lynch, R. Sturani, and P. Graff. Use of gravitational waves to probe the formation

channels of compact binaries. *Classical and Quantum Gravity*, 34(3):03LT01, Jan. 2017b. doi: 10.1088/1361-6382/aa552e. URL https://doi.org/10.1088/1361-6382/aa552e.

R. Voss and T. M. Tauris. Galactic distribution of merging neutron stars and black holes–prospects for short gamma-ray burst progenitors and LIGO/Virgo. *Monthly Notices of the Royal Astronomical Society*, 342(4):1169–1184, 2003.

K. W. K. Wong, K. Breivik, K. Kremer, and T. Callister. Joint constraints on the field-cluster mixing fraction, common envelope efficiency, and globular cluster radii from a population of binary hole mergers via deep learning. *Phys. Rev. D*, 103:083021, Apr 2021. doi: 10.1103/PhysRevD.103.083021. URL https://link.aps.org/doi/10.1103/PhysRevD.103.083021.

D. Wysocki, D. Gerosa, R. O'Shaughnessy, K. Belczynski, W. Gladysz, E. Berti, M. Kesden, and D. E. Holz. Explaining LIGO's observations via isolated binary evolution with natal kicks. *Phys. Rev. D*, 97:043014, Feb 2018. doi: 10.1103/PhysRevD.97.043014. URL https://link.aps.org/doi/10.1103/PhysRevD.97.043014.

D. Wysocki, J. Lange, and R. O'Shaughnessy. Reconstructing phenomenological distributions of compact binaries via gravitational wave observations. *Phys. Rev. D*, 100:043012, Aug 2019. doi: 10.1103/PhysRevD.100.043012. URL https://link.aps.org/doi/10.1103/PhysRevD.100.043012.

B. Zackay, L. Dai, and T. Venumadhav. Relative binning and fast likelihood evaluation for gravitational wave parameter estimation, 2018.

B. Zackay, L. Dai, T. Venumadhav, J. Roulet, and M. Zaldarriaga. Detecting Gravitational Waves With Disparate Detector Responses: Two New Binary Black Hole Mergers, 2019a.

B. Zackay, T. Venumadhav, L. Dai, J. Roulet, and M. Zaldarriaga. Highly spinning and aligned binary black hole merger in the Advanced LIGO first observing run. *Phys. Rev. D*, 100:023007, Jul 2019b. doi: 10.1103/PhysRevD.100.023007. URL https://link.aps.org/doi/10.1103/PhysRevD.100.023007.

B. Zackay, T. Venumadhav, J. Roulet, L. Dai, and M. Zaldarriaga. Detecting gravitational waves in data with non-Gaussian noise, 2019c.

M. Zaldarriaga, D. Kushnir, and J. A. Kollmeier. The expected spins of gravitational wave sources with isolated field binary progenitors. *Monthly Notices of the Royal Astronomical Society*, 473(3):

4174–4178, 10 2017. ISSN 0035-8711. doi: 10.1093/mnras/stx2577. URL https://doi.org/10.1093/mnras/stx2577.

M. Zevin, C. Pankow, C. L. Rodriguez, L. Sampson, E. Chase, V. Kalogera, and F. A. Rasio. Constraining formation models of binary black holes with gravitational-wave observations. *The Astrophysical Journal*, 846(1):82, sep 2017. doi: 10.3847/1538-4357/aa8408. URL https://doi.org/10.3847/1538-4357/aa8408.

M. Zevin, S. S. Bavera, C. P. L. Berry, V. Kalogera, T. Fragos, P. Marchant, C. L. Rodriguez, F. Antonini, D. E. Holz, and C. Pankow. One channel to rule them all? constraining the origins of binary black holes using multiple formation pathways. *The Astrophysical Journal*, 910(2):152, Apr. 2021. doi: 10.3847/1538-4357/abe40e. URL https://doi.org/10.3847/1538-4357/abe40e.

B. M. Ziosi, M. Mapelli, M. Branchesi, and G. Tormen. Dynamics of stellar black holes in young star clusters with different metallicities–II. Black hole–black hole binaries. *Monthly Notices of the Royal Astronomical Society*, 441(4):3703–3717, 2014.

S. F. P. Zwart and S. L. McMillan. Black hole mergers in the universe. *The Astrophysical Journal Letters*, 528(1):L17, 1999.